Jolivet, G., & Turon, H. (2018). Consumer Search Costs and Preferences on the Internet. *Review of Economic Studies*, *86*(3), 1258-1300. [rdy023]. https://doi.org/10.1093/restud/rdy023

Peer reviewed version

Link to published version (if available):
10.1093/restud/rdy023

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Oxford University Press at https://academic.oup.com/restud/article-abstract/86/3/1258/4993838?redirectedFrom=fulltext . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Consumer Search Costs and Preferences on the Internet*

Grégory Jolivet †
University of Bristol

Hélène Turon‡
University of Bristol

December 2017

Abstract

We conduct an empirical analysis of consumer preferences and search costs on an Internet platform. Using data from a major French platform (PriceMinister), we show descriptive evidence of substantial price dispersion among adverts for the same product, of consumers often not choosing the cheapest advert and sometimes choosing an advert dominated in price and non-price characteristics by another available advert. We consider a sequential search model where consumers sample adverts in an endogenous order based on their preferences and search costs. We show that the optimal search-and-purchase strategy can be characterised by a set of inequalities which can feasibly be tested on transaction and advert data. This allows us to estimate, for each transaction, a set of preference and search cost parameter values, thus allowing for flexible consumer heterogeneity in preferences and search costs. The estimated model can then describe a wide range of consumer search and purchase behaviours. We find that the model can explain almost all transactions in the data and requires non-zero preferences and search costs for at least 50% and 22% (respectively) of observations. We also find evidence of heterogenous and sometimes substantial search costs.

Keywords: Consumer Search, Individual Heterogeneity, Price Dispersion, Internet.

JEL Codes: C13, D12, D81, D83, L13

*We are deeply grateful to PriceMinister.com for giving us access to their data. We thank the editor and four anonymous referees for helping us improve our work. We also thank Ian Crawford, José Moraga-González, David Pacini and Sarit Weisburd for giving us very helpful feedback, as well as audiences at the 2014 SaM, 2014 EARIE, 2016 CEPR-JIE conferences and at various seminars for their questions and comments.

†Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK. E-mail: gregory.jolivet@bris.ac.uk.

‡Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK. E-mail: helene.turon-lacarrieu@bris.ac.uk.

1 Introduction

The advent of e-commerce, in particular Internet platforms, was initially presumed to increase competition and thus decrease price dispersion, since it allowed the gathering of information at little physical and time cost. However, casual observation of trading websites as well as the emergence of rich datasets documenting the variance of prices among adverts and transactions convey a compelling message: price dispersion remains and can be substantial (see e.g. Baye et al., 2004). Potential sources of dispersion have been investigated by the recent economic literature. First, differences in non-price characteristics across sellers or adverts can generate price dispersion even for a specific product. Second, acquiring or processing information about all relevant characteristics may still be costly for consumers and these search costs may reduce the scope for the "law of one price" to prevail. Accordingly, the consumer search literature has been at the forefront of the analysis of Internet markets. A recent paper by Anderson and Renault (2016) provides a detailed survey of this literature, stressing that "appropriate theoretical frameworks should involve sufficient heterogeneity among agents on both sides of the market" and also explaining that "the analysis of ordered search constitutes an essential ingredient for modelling recent search environments". The latter point follows from the fact that consumers may search across alternatives in a specific, possibly endogenous, order.

In this paper, we provide empirical evidence on the importance of consumer preferences and search costs based on a new structural estimation approach. We consider the (partial-equilibrium) model of Weitzman (1979), where consumers search sequentially for alternatives in an endogenous order and we show that the optimal search-and-purchase strategy can be characterised by a simple set of inequalities. These inequalities lend themselves to a feasible set-estimation approach inspired by the revealed preferences literature. One strength of our approach is that we allow for almost unconstrained heterogeneity in preferences and search costs. This, combined with the endogenous sampling strategies predicted by the model, leads our structural model to account for a wide range of consumer search behaviours and rationalise nearly all transactions observed on the platform under study. Our estimation approach and the model we use thus makes our analysis contribute to the two main features of the consumer search literature put forward by Anderson and Renault (2016).

We estimate the model on an administrative record of all transactions and all adverts on the website priceminister.com, which is one of the main e-commerce platforms in France. For each transaction, we have information on all the adverts that were available at that particular date for the very same product, identified by a barcode. These are the adverts that were on the computer screen when the consumer searched for this specific product on this website. In this paper, we will focus on CD transactions. Adverts vary in price but also in other characteristics such as the condition of the item, the seller's reputation etc. We show substantial price dispersion among adverts and among transactions for the same product. We also find that consumers often do not buy the cheapest available advert and sometimes even choose an advert that is clearly dominated (in price and in non-price characteristics) by another available advert. These stylised facts motivate our structural analysis which focuses on two objects: consumers' taste for non-price characteristics and search costs.

Our structural analysis builds on a sequential model of directed search (on prices) *à la* Weitzman (1979).

We now briefly describe this framework to fix ideas. Assume that a consumer wants to buy a specific product and is presented with a finite number of adverts. Consistently with the design of the PriceMinister website, we assume that prices are observable at no cost and may thus be used by the consumer to determine his optimal search order, but that consumers must pay a search cost to examine an advert's characteristics and compute the corresponding utility.[1] The consumer samples adverts one at a time and incurs the same search cost for each draw. The consumer chooses in which order he samples adverts, recall is allowed and the search can stop before all adverts have been sampled.

The optimal search and purchase strategy has been derived by Weitzman (1979) and depends on the consumer's beliefs about advert characteristics given price and on the consumer's search cost and marginal willingness to pay (MWP hereafter) for these characteristics. Combining this search-and-purchase rule with heterogeneity in consumer MWPs and search costs leads to an important feature of our analysis, which is that it can describe a wide range of sampling patterns. In particular consumers do not necessarily search in ascending price order and/or do not necessarily buy the last item sampled. Some consumers will sample very few adverts, others will almost exhaust all the offers. These different patterns will arise from heterogeneity in search costs and in consumer preferences. Hence, a consumer's choice set is formed endogenously, depending on his (individual-specific) MWP and search cost. Whilst the optimal stopping rule is often incorporated in consumer search models, the analysis of the optimal search order as a direct consequence of the consumer's preferences and search costs within a structural estimation is one of the main innovations of our paper.

To estimate the model we derive two analytical results. First, we show that the optimal search and purchase rule derived by Weitzman (1979) is equivalent to a set of tractable inequalities. Second, we derive a simple characterisation of the set of parameters consistent with each transaction. Estimation proceeds in three steps. The first step consists in projecting advert non-price characteristics onto one "hedonic" scalar. The second step estimates consumers' beliefs about the distribution of this hedonic characteristic conditionally on advert price, using the observed joint distribution of advert prices and characteristics. The last step is close in spirit to the revealed preference literature (see Blow et al. (2008), Cherchye et al. (2009), Cosaert and Demuynck (2014)) and uses the conditions derived from our theoretical analysis to find, for each transaction, which values of the MWP and search cost parameters are consistent with the fact that the bought advert was chosen over the alternative adverts. We do not impose any restriction on the shape of consumer heterogeneity with respect to search costs or to the MWP for the advert "hedonic" scalar.[2] Note also that we are not excluding the case where the consumer only cares about prices (i.e. his MWP is zero) and/or faces no search costs (and can sample all adverts for free). Our approach will then produce bounds on the distributions of MWPs and search costs which will be informative to assess the importance of search costs and consumer preferences in our data. We are, to the best of our knowledge, the first to use this empirical approach in the consumer search literature.

Our benchmark specification can explain 95% of the CD transactions observed on the website in a specific

[1]Since consumers can use prices to direct their search towards specific adverts, we will say that our model is a directed sequential search model. Search *à la* Weitzman (1979) could also be referred to as ordered search but this could also mean that the search order is imposed whereas in our case, as in the original Weitzman (1979) model, it is chosen by the consumer.

[2]We do however need to specify a functional form for the individual utility function, thus departing from the revealed preference literature.

quarter in 2007, which is our benchmark estimation sample. As mentioned above, many transactions are such that the advert sold is dominated by an alternative advert in price and hedonic characteristics. A hedonic perfect information model cannot explain these transactions (unless consumers only care about prices and the two adverts are at the same price), whereas our model is able to rationalise 76% of these transactions with reasonable values of the MWP and search costs. Since the cheapest advert is bought in slightly less than 50% of transactions, a strictly positive MWP is needed to explain half of the transactions in our data. Importantly, we find that strictly positive search costs play a substantial role on the Internet platform we study. Indeed, 22% of the transactions we explain require a search cost strictly larger than zero and are thus inconsistent with a perfect-information model.

To investigate the degree of heterogeneity in preferences and search costs in our set-estimation approach, we borrow from a recent article by Stoye (2010) and compute, for each of these parameters, the minimal variance compatible with our estimated bounds. Whilst the minimum-variance distribution of MWP's displays little dispersion, we find evidence of heterogeneity in search costs.

Our estimated model illustrates a varied set of search patterns. For example, when setting individual search costs to the minimum value for each explained transaction, we find that consumers who face strictly positive search costs buy the first advert that they sample 63% of the time. Besides we find that for 13% of transactions with positive search costs, the consumer has carried on sampling after finding the advert that he eventually buys. The model estimates also allow us to quantify the utility cost incurred by consumers who face strictly positive search costs, and whether this cost comes mainly from searching (sampling cost) or from the fact that consumers do not always buy the best available advert.

Our benchmark model estimation relies on observed advert characteristics and rules out that unobserved attributes may affect the consumer's decision. We show in an extension that our approach can account for unobserved characteristics, at the cost of making a parametric assumption on their distribution. Our estimation strategy can still be adopted and produces results which are in line with those we find in our benchmark estimation.

We end this introduction with a review of the consumer search literature that this paper contributes to. Early empirical works include Hortaçsu and Syverson (2004) who consider that investors search sequentially for funds and have homogenous preferences for fund attributes and heterogenous search costs. In their framework, the link between a fund's sampling probability and the investor's preferences is not modeled. Hong and Shum (2006) estimate a search cost distribution for consumers searching for the best price (no other product attributes) for academic textbooks, considering both a sequential and a non-sequential search model. An alternative parametric approach exploiting the equilibrium conditions of the non-sequential search model has been suggested by Moraga-González and Wildenbeest (2008) and a semi-parametric approach has been developed by Moraga-González et al. (2013). De Los Santos et al. (2012) use data on the search and purchase of books across different Internet websites and reject sequential search on the grounds that consumers do not necessarily buy from the last store visited. As we argue in our paper, allowing for directed sequential search instead of random search makes our sequential search model consistent with this feature. Two recent articles have estimated search costs in models where consumers learn about the distribution of utilities through searching: Koulayev (2014) and De Los Santos et al. (2017). In these papers the order in which adverts are sampled is either random or imposed, but not linked to consumer preferences. In contrast, our approach allows for endogenous sampling order but rules out learning.

Sequential directed search models building on Weitzman (1979) have recently been estimated, mostly using Internet data. The first paper we are aware of is Kim et al. (2010) and more recent contributions include Kim et al. (2016), Ursu (2017), Honka and Chintagunta (2016) and Chen and Yao (2016).[3] Our paper departs from these existing studies on at least three dimensions. First, we do not not require data on search behaviour to (set) identify search costs and preferences, we only use choice and offer data. Our approach can thus be taken to a large set of applications whereas search data (i.e. data on which options consumers "click" on) are still relatively rare. Secondly, in most of these papers, the advert attribute that is unobserved prior to sampling is modelled as an unobserved (to the econometrician) taste shock, the distribution of which is parameterised and independent of the observed characteristics driving consumer search. The two exceptions we are aware of are Kim et al. (2016), where the unobserved attribute's variance may depend on ex-ante observable characteristics and Chen and Yao (2016), where consumers can sort adverts along one dimension and then infer on the advert characteristics from its position on the computer screen. In our paper, we assume that prices are observed ex ante but that consumers must pay a search cost to find out the advert "hedonic" index, which is allowed to be related to the price through a flexibly estimated function. We think that this is closer to Weitzman (1979)'s model as prices can be seen as a signal that carries information on the distribution of characteristics. The third difference between our paper and previous studies is that we allow for a very flexible modelling of consumer heterogeneity in both preferences and search costs and are thus able to capture a wide range of search behaviours. Indeed, some transactions will be explained by a perfect-information model (no search cost), others by a model where consumers search in increasing price order and others where search order is non monotonic in price. Within our approach, the search costs and MWPs we find for a given transaction do not influence those found for other observations through some parametric distribution of either of these two parameters. We believe this allows us to give robust empirical evidence on search costs and preferences on the Internet platform we study.

Since our main estimation targets are demand-side structural objects, namely consumer preferences and search costs, we retain a partial equilibrium analysis and focus on developing a flexible estimation approach whilst allowing for heterogeneity and elaborate search strategies. Naturally, our results trigger questions related to the price-setting behaviour of sellers in view of these heterogenous search costs and preferences on the consumers' side. We will however leave these questions for future research and refer here to the relevant literature for an equilibrium analysis. The theoretical literature on consumer search models is vast, starting with Stigler (1961) and continuing with important contributions in the 1980s (a non-exhaustive list includes Burdett and Judd, 1983, Wolinsky, 1986, Stahl II, 1989) aiming at overcoming the Diamond (1971) paradox. More recent contributions include Anderson and Renault (1999) who study consumer search product differentiation, Janssen and Moraga-González (2004) who analyse firm behaviour with oligopolistic competition and non-sequential search, or Moraga-González and Petrikaite (2013) who derive an equilibrium

[3]Directed sequential search models have also been studied by the decision science literature, see for instance Analytis et al. (2014).

sequential search model where consumers can direct their search towards merging firms depending on their expectations over price. On a more empirical branch of the literature, Zhou (2011) presents an (exogenously) ordered search model in which firms visited late in the search process enjoy some monopoly power since consumers visiting them do so when they have a low valuation of the products offered by firms already visited.[4] Another recent paper by Dinerstein et al. (2014) uses rich data on eBay to estimate an equilibrium non-sequential search model with homogenous consumers and to simulate the effects of changes in the platform's design.

Our paper is organised as follows. Section 2 details the model and derives the conditions used to identify the sets of preference and search cost parameters. Section 3 presents our data and gives empirical evidence on price dispersion and consumer search and purchase behaviours on an Internet platform. Section 4 details the three steps of our empirical strategy and section 5 shows the estimation results on fit, search costs, preferences and heterogeneity. Section 6 uses the model estimates to describe consumer search behaviour. An extended version of model is derived and estimated in section 7 and section 8 concludes.

2 The model

2.1 Consumers' search and purchase decision

Consider a buyer who wishes to purchase one unit of a specific product (for instance a given CD) on an Internet platform. Let $J \geq 1$ be the number of adverts for this product that are currently posted on the platform. Each advert $j \in \{1, ..., J\}$ consists of a price p_j and a vector of characteristics x_j. In our application, x will contain the seller's reputation index, its size (i.e. the number of transactions carried out on the website to date), its status (professional or not) and the condition of the good being sold. We assume that the consumer's outside option, i.e. the utility of not buying anything, is very low so that one advert is always bought.[5]

Preferences. The utility of buying an advert (p, x) is given by:

$$u(p, x, \gamma) = \gamma x - p. \tag{1}$$

The parameter γ, which has the same dimension as x, is the consumer's marginal willingness to pay (MWP) for the advert's characteristics and summarises the consumer's preferences for these characteristics. We allow consumers to have heterogenous preferences so γ may vary across individuals. For notational convenience, we will sometimes write the utility offered by advert j as u_j.[6]

[4] In contrast with our analysis, Zhou (2011) focuses on consumers sampling adverts by increasing order of price.

[5] It would be straightforward to add an outside option value to the following partial-equilibrium model. However, we will be using data on transactions and will not observe consumers who just visit the website and do not buy anything (i.e. we do not have search data). We will thus not be able to account for the selection into the set of buyers, which may depend on preferences and search costs. Hence, we maintain the assumption that the outside option is very low (more precisely low enough for search and purchase to take place) and we will conduct our estimation while keeping in mind that our results hold for the population of consumers how actually buy a product on the website.

[6] Moreover, since we will eventually project all advert non-price characteristics onto a scalar (see section 2.4), we write the product γx as if it were a product of scalars.

Search. In order to acquire information on adverts' relevant characteristics, consumers may need to engage in a costly search process. We assume that consumers search sequentially, with possibility of recall. This means that they sample advert one at a time and can always buy any advert previously sampled.[7]

Drawing an advert incurs a search cost $s \geq 0$ which, like the preference parameter γ, is allowed to vary between consumers. Drawing an advert means being able to observe all the advert characteristics and thus finding out the level of utility it offers. The search cost s can be thought of in a number of ways, such as a cost of looking at the advert characteristics or a cost of processing the information given by the new advert. This search cost is assumed to be constant across draws. Note that we do not rule out the case $s = 0$ i.e. the perfect information model whereby a consumer simply buys the advert offering the highest utility.

We assume that consumers can see all the available advert prices instantly and at no cost but have to pay the search cost s to observe a given advert's characteristics x. This follows from the design of the website used in our empirical application. When consumers are looking at adverts for a given product, they first see all adverts ranked by increasing order of price. The price is shown in a larger font than other characteristics (such as seller reputation etc.). We thus think that it is realistic to assume that collecting information on prices is costless for consumers but that they must pay a utility cost to gather additional information on adverts as these details are less visible and not ranked by default. Consumers can use advert prices to direct their search.

Note that in this directed search model, the perfect-information case obtains not only when $s = 0$ but also when $\gamma = 0$. Indeed, if a consumer only cares about prices and if information about advert prices is available at no (search) cost, this consumer will look at all the advert prices and buy the cheapest advert, as if there were no search frictions.

Beliefs and advert distribution. The consumer believes that the J adverts presented to him are independent draws from a joint distribution of prices and characteristics (P, X), denoted $F_{P,X}$.[8] We assume that consumers' beliefs are homogeneous and remain the same during the search process, i.e. we rule out learning.[9] This will allow us to derive a simple optimal search strategy for consumers. We also need to consider the conditional distribution $F_{X|P}$, which is what consumers believe to be the cdf of X for a given price P.

Let $F^0_{P,X}$ denote the cdf of prices and characteristics in the population of all adverts actually posted on the platform (for a given product category) in a given time window. This distribution follows from sellers' pricing strategies. For instance, sellers may differ with respect to their characteristics x and, given their value of x and the distribution of consumers' preferences and search costs, set prices that maximise their expected profit. The resulting distribution, $F^0_{P|X}$, combined with the distribution of seller characteristics F^0_X leads to the observed distribution of prices and characteristics $F^0_{P,X}$. In this paper, since we restrict our

[7] The alternative is a non-sequential search environment whereby consumers decide on an optimal number of draws ex-ante (see e.g. Morgan and Manning, 1985 for an analysis of optimality). We believe the sequential search to be more realistic to model search on an Internet platform. A recent article by De Los Santos et al. (2012) shows that the behaviour of consumers looking for books *across* websites is not consistent with a sequential random search model, as consumers sometimes buy from a previously visited website. In this paper however, we will consider a *directed* search model so sequential search will be consistent with consumers retracing their steps.

[8] We use capital letters for random variables and small letters for their realisations.

[9] See e.g. Koulayev, 2014 for a recent analysis of consumer search with learning.

analysis to a partial equilibrium, we will take $F^0_{P,X}$, which we can directly observe in the data, as given.

A natural way to anchor consumers' beliefs would be to assume that $F_{P,X} = F^0_{P,X}$. In a full-equilibrium setting, this means that consumers' beliefs are consistent with sellers' pricing strategies. From an empirical perspective, this assumption allows us to estimate consumers' beliefs as the observed distribution of prices and characteristics in the population of adverts. However, as we will discuss in detail in Section 4.2, one may impose further restrictions on the beliefs in order to limit the level of sophistication in consumers' predictions.

2.2 Consumers' search and purchase decision

We already know that a consumer with $s = 0$ will buy the advert offering the highest utility and that a consumer with $\gamma = 0$ will buy the cheapest advert. We now present the optimal search and purchase strategy used by consumers for whom $s > 0$ and $\gamma > 0$. This strategy has been characterised in an influential article by Weitzman (1979). We will first present this result in the context of our model and then show how it can be used to identify consumers' preference and search cost parameters. We give a more formal presentation of the consumers' sequential search problem in Appendix A.

A first step is to define the following quantity, which will drive the individual consumer's search strategy: the "reservation utility" of an individual with preference parameters (s, γ), denoted $r(p, s, \gamma)$, which represents the utility level at which this individual is indifferent between sampling a new item priced p, thus incurring a search cost s, and not sampling it, thus staying with a utility of $r(p, s, \gamma)$. This threshold utility satisfies the following condition:

$$s \quad = \quad E_{X|P=p}\Big[\big(u(p, X, \gamma) - r(p, s, \gamma)\big) \cdot \mathbf{1}\left\{u(p, X, \gamma) > r(p, s, \gamma)\right\}\Big], \tag{2}$$

where the expectation is taken over the distribution of advert characteristics X given that the price is equal to p, and $\mathbf{1}\{\cdot\}$ is the indicator function. In other words, the expected gain over $r(p, s, \gamma)$ of sampling this item is equal to the search cost s.[10]

It is apparent from (2) that this "reservation utility" depends on the price p, on the individual parameters γ and s but also on consumers' beliefs regarding the joint distribution of X and P. We will sometimes denote the reservation utility offered by advert j simply as r_j, instead of $r(p_j, s, \gamma)$. Note that all consumers share the same beliefs but are heterogeneous with respect to (s, γ). This model rationalises the search order and this order will vary across consumers as the sequence of reservation utilities is individual specific.

The optimal sequential search and purchase strategy, as derived by Weitzman (1979), is the following:

- The consumer draws adverts in decreasing order of reservation utilities r.
- The consumer stops searching when the highest utility he has drawn is larger than the largest reservation utility among adverts not yet drawn.
- When stopping the consumer buys the advert with the highest utility among the ones he has sampled.

[10] Equation (2) defines one and only one reservation utility as the search cost s is positive and the right-hand side of (2) is a continuous and strictly decreasing function of r with a range of $(0, +\infty)$.

Ties are assumed to be resolved in the following way. If several adverts have the same reservation utility r, consumers sample them in a random order. If several adverts that have been drawn offer the same maximum level of utility, the consumer chooses one randomly. When indifferent between stopping his search and sampling another advert, the consumer stops searching.

2.3 Identification of preferences and search costs

In the spirit of the revealed preference literature (see Blow et al., 2008), we now identify sets of parameters that are consistent with the choices observed in the data. Consider a transaction where advert i is sold. We may also refer to this transaction as transaction i.[11] From now on, for each transaction, all quantities (p, x, u, r) will be indexed by i if they refer to the advert bought, and by j if they refer to an advert which was available (i.e. on the screen) but was not bought.

We derive necessary and sufficient conditions for a pair (s, γ) to be consistent with the fact that i was bought while advert j was also available but not chosen. These conditions will characterise a set S_{ij}. We can then define the set S_i of parameters consistent with transaction i as the intersection of all the sets S_{ij} for all available adverts for this transaction.[12]

First, we assess whether a transaction can be explained by a perfect information model:

$$(s = 0, \gamma) \in S_i \quad \Leftrightarrow \quad \gamma x_i - p_i \geq \max_j \{\gamma x_j - p_j\}, \tag{3}$$

$$(s, \gamma = 0) \in S_i \quad \Leftrightarrow \quad p_i \leq \min_j \{p_j\}. \tag{4}$$

In words, consumers with no search costs should buy the best advert and consumers who only care about prices should buy the cheapest advert (since information on prices is available for free).

We now turn to the case where both the search cost s and the preference parameter γ are strictly positive. The identified set S_{ij} is characterised as follows:

Proposition 2.1. *Let $s > 0$ and $\gamma > 0$. Then $(s, \gamma) \notin S_{ij}$ if and only if*

$$\Big(u_i \geq r_i \quad \textit{and} \quad r_j > r_i \quad \textit{and} \quad u_j \geq r_i\Big) \tag{5}$$

$$\textit{or}$$

$$\Big(u_i < r_i \quad \textit{and} \quad r_j > u_i \quad \textit{and} \quad u_j > u_i\Big). \tag{6}$$

The proof is in Appendix B. This gives us a simple accept/reject rule for each value of (s, γ) to rationalise the transaction i being observed while advert j was present on the screen. Our characterisation of S_{ij} thus follows from simple inequalities. For each transaction i and each value of (s, γ), we just need to compute the instantaneous and reservation utilities and check whether (5) or (6) holds for any advert j. If this is not the case, this parameter value rationalises the transaction. The problem with this approach is that reservation utilities r are only implicitly defined by an equation that involves both preference and search

[11] In a slight abuse of language, we refer to the distribution of parameters across transactions as the distribution of parameters across purchasing consumers. This is only valid if all consumers buy exactly once, and in the absence of information of consumers' identities in the data, we are not in a position to confirm this.

[12] If the data allowed us to identify consumers, we would be able to narrow the set even further by considering the intersection of all the S_i's pertaining to purchases made by a consumer.

cost parameters, see equation (2). In the next section, we show how we make our characterisation of the identified sets far more tractable.

2.4 The case of a scalar hedonic index

So far, we have considered a general case where the non-price characteristics of adverts consisted of a vector x. In this section, we assume that x is a scalar and that it is valued positively by all consumers so that $\gamma \geq 0$. γ now represents the individual's marginal willingness to pay (MWP) for the hedonic index x. This assumption and the results from the previous section lead a more tractable characterisation of the sets of identified parameters.

In practice, the scalar index x will be obtained as a function of advert characteristics (in our case, x will be a linear index of seller's reputation, product condition, etc.), which is assumed to be the same for all consumers. This approach still allows for heterogeneity in preferences but slightly restricts it in that, under this assumption, consumers have homogeneous marginal rates of substitution between two non-price advert characteristics. However the MWP γ for the hedonic index x varies across consumers, indeed there is no restriction on the distribution of the scalar γ other than it being positive or zero. This is not restrictive as, in our data, we can find an advert characteristic for which the MWP is unlikely to be strictly negative (for instance seller reputation or product condition). We will show in detail in section 4.1 how we construct a structural projection of advert characteristics onto the scalar hedonic index x.

We now introduce the following function:

$$\psi_p(x) \quad = \quad E_{X|P=p}\left[(X-x)\cdot \mathbf{1}\left\{X>x\right\}\right]. \tag{7}$$

$\psi_p(x)$ reflects the expected gain over x (the scalar hedonic index) when an item of price p is sampled. An important feature of this function, which will be very useful for estimation, is that it does not depend on (s, γ).

The function ψ_p is differentiable and strictly decreasing in x on the support of X given price p. We can thus define its inverse ψ_p^{-1}. Note also that ψ characterises the consumer's beliefs and that the condition $\psi_{p'} \geq \psi_p$ is equivalent to second-order stochastic dominance of $F_{X|P=p'}(\cdot)$ over $F_{X|P=p}(\cdot)$ (see Hadar and Russell (1969)). Hence, if consumers believe that the distribution of X improves with price, this is reflected by the fact that ψ_p increases with p.

Importantly, we can use the definitions in equations (2) and (7), to derive another expression for $r(p, s, \gamma)$ which conveniently mirrors that of utility:

$$r(p, s, \gamma) = \gamma \cdot \psi_p^{-1}\left(\frac{s}{\gamma}\right) - p. \tag{8}$$

We can now give a tractable characterisation of the identified set S_{ij}. First, we use conditions (3) or (4) for the cases where $s = 0$ or $\gamma = 0$ respectively. With a scalar hedonic index, these conditions are easier to interpret. Indeed, let us say that advert j (alternative) is 'better' than advert i (bought) if j is both cheaper ($p_j \leq p_j$) and of better quality ($x_j \geq x_i$) than i with at least one of these inequalities being strict. When there exists an advert j that is 'better' than the bought advert i then the transaction i cannot be explained by a perfect-information model ($s = 0$) unless $p_i = p_j$ in which case the transaction can be explained by $s = \gamma = 0$.

On the other hand, the absence of a 'better' alternative advert does not necessarily mean that a transaction can be explained without search costs. For example, consider the case where advert i is sold whilst adverts j and k were available, advert i is cheaper and offers less quality than j and is more expensive and offers better quality than k. Hence, neither j nor k is 'better' than i. Assume further that $\frac{p_j - p_i}{x_j - x_i} < \frac{p_i - p_k}{x_i - x_k}$. In this case a pair $(s = 0, \gamma)$ cannot be in S_i since the value of γ would have to be lower than $\frac{p_j - p_i}{x_j - x_i}$ (the MWP has to be low enough to rationalise i being preferred to j) and larger than $\frac{p_i - p_k}{x_i - x_k}$ (the MWP has to be high enough to rationalise i being preferred to k). Hence a perfect-information model cannot explain this transaction.

If s and γ are strictly positive, we plug the expressions of the utility and reservation utility, (1) and (8), into the inequalities (5) and (6) to obtain the following characterisation:

Proposition 2.2. *Let x be a scalar, $s > 0$ and $\gamma > 0$. Then $(s, \gamma) \notin S_{ij}$ if and only if*

$$\left(\frac{s}{\gamma} \geq \psi_{p_i}(x_i) \quad \textit{and} \quad \gamma \left[\psi_{p_j}^{-1}\left(\frac{s}{\gamma}\right) - \psi_{p_i}^{-1}\left(\frac{s}{\gamma}\right) \right] > p_j - p_i \quad \textit{and} \quad \gamma \left[x_j - \psi_{p_i}^{-1}\left(\frac{s}{\gamma}\right) \right] \geq p_j - p_i \right) \tag{9}$$

or

$$\left(\frac{s}{\gamma} < \psi_{p_i}(x_i) \quad \textit{and} \quad \gamma \left[\psi_{p_j}^{-1}\left(\frac{s}{\gamma}\right) - x_i \right] > p_j - p_i \quad \textit{and} \quad \gamma (x_j - x_i) > p_j - p_i \right) \tag{10}$$

The main advantage of (9)-(10) compared to (5)-(6) is that the conditions are now simple plug-in functions of the parameter values (s, γ). Once we have an estimate of the ψ_p^{-1} functions for each price (and this can be done without looking at consumers' choices), finding the set of parameters consistent with a given transaction can be done by a simple grid-search method, using (9)-(10) as a pass/reject criterion.

3 Data and descriptive statistics

3.1 The PriceMinister website

We use data from PriceMinister, a French company organising online trading of new and second-hand products between buyers and professional or non-professional sellers. We will focus on the company's French website www.priceminister.com. PriceMinister is one of the largest e-commerce websites in France with 11 million registered users and over 120 millions products for sale in 2010 (the site opened in 2001).[13] Whilst many different goods can be bought from the website (books, television sets, shoes, computers), we will focus on CDs and, in a robustness check, on DVDs. 'Cultural' goods (books, CDs, video games and DVDs) represented the vast majority of transactions during our observation period.

The website is a platform where sellers, professional (registered businesses) or non professional (private individuals), can post adverts for goods, which can be used or new.[14] When a potential buyer searches for

[13]PriceMinister was ranked first among e-commerce websites in terms of ratings in a survey conducted by Mediamétrie in March 2010. The other main e-commerce websites in France were Amazon, eBay and Fnac.

[14]PriceMinister does not charge a sign-on fee, and posting an advert is free of charge. However, for each completed transaction, sellers have to pay a variable fee to PriceMinister. The fee scale is posted on the priceminister.com website. Only professional sellers are allowed to advertise items as 'new'.

a specific item, the website returns a page of available adverts. These include the price (adverts are sorted by increasing prices by default), the condition of the item: new or used ('as new', 'very good', 'good'), the seller's status (professional or not), reputation (to be defined shortly) and size (the number of transactions completed by the seller).[15]

In this paper, we focus on the consumer's search behaviour when faced with a page of adverts for a specific product. We do not model how the consumer has decided to reach this page and this website. Since our data relate to transactions, we know that the consumer must have reached the page of adverts for this product before he made his purchasing decision, and we carry out our analysis conditionally on this.

A seller's reputation is the average of feedbacks received since the creation of the seller's account. To explain the feedback mechanism, we now describe how transactions take place on the website. When a buyer purchases a given product from a given seller, the buyer's payment is made to PriceMinister in the first instance. At this point the seller is informed that a buyer has chosen her product and ships the item to the buyer. Once the buyer has received the product, he is prompted to give his feedback on the transaction through the website. PriceMinister then closes the transaction and pays the seller.[16] The buyer's feedback consists of a grade, or rating, which is any integer between 1 (very disappointed) and 5 (very satisfied). The default provided on the feedback form is equal to 5.[17] The seller's reputation as posted on the website is the average of the feedbacks received for all completed transactions, rounded to the nearest first decimal. A seller's size at a given date is then the number of transactions that she has completed so far.

We should mention that PriceMinister differs from other e-commerce websites that are studied in the economic literature with respect to several features that are important for our analysis. First, PriceMinister itself does not sell any products: it is a platform (unlike, e.g., Amazon). Hence consumers may not direct their search towards the seller that also operates the platform. Secondly, prices are posted by sellers, there are no auctions (unlike eBay).[18]

3.2 The data

We were given two administrative datasets by PriceMinister: one with all the transactions between 2001 and 2008, and one with all the adverts posted between 2001 and 2008. In this paper, we will focus on transactions of CDs which took place in the third quarter of 2007. Unless otherwise mentioned, all the following descriptive statistics and estimation results will be produced for this selected sample. Results for another time period or product category (DVD) are in Appendix G.

For each transaction or advert, we observe the price, product and seller ID (not the buyer ID), and all the characteristics mentioned above (product condition, seller's status, reputation and size). We can thus construct a dataset where, for each transaction, we observe all the adverts that were on the screen for the

[15] The advert also shows the seller's name, country and the different shipping options. In this version of the paper, we do not include sellers' country in the characteristics vector x because it is France for the overwhelming majority of sellers in our sample. We will discuss shipping options and costs later on in Section 3.2.

[16] When a buyer files a complaint, PriceMinister investigates and puts the payment on hold. If the buyer does not contact PriceMinister within 6 weeks, he is sent a reminder e-mail. If he does not respond, PriceMinister closes the transaction and pays the seller.

[17] The fact that buyers must give feedback in order to validate the transaction ensures a high feedback rate (above 90% for transactions with individual sellers).

[18] In recent years, buyers may be offered the option to negotiate the price but this option was introduced at the end of our observation period and, at the time, very rarely used.

same specific product when the consumer made his choice.[19] Note that products are precisely identified on the website (for instance by their barcode). Our empirical analysis allows the consumer to value not only the item purchased but the bundle of the item and the service provided by the seller (response rate and speed, delivery time).

Our benchmark analysis will assume that no characteristics other than the price and the four ones included in x are either observed or valued by the buyer. In particular there is no advert heterogeneity that may be observed by the buyer and unobserved by the econometrician. We end this section with a discussion of two potential violations of this assumption. To fix ideas, we show and comment on a screen shot (in French) of an advert page for a specific CD in November 2007 in Appendix D.

For the price variable, we use the advert price net of shipping costs. On PriceMinister, sellers cannot differentiate themselves with respect to shipping costs. In any transaction, the choice of shipping mode (essentially standard, registered mail or fast delivery) is up to the buyer, subject to a fixed shipping cost menu imposed by PriceMinister. Specifically, the buyer chooses a particular shipping option at the time of purchase and the corresponding fee on the shipping cost menu is added to the bill and transferred to the seller by PriceMinister. It is then up to the seller to minimise its costs, subject of course to complying with the buyer's specific choice of shipping mode. Sellers could still differentiate themselves by offering an additional type of shipment that may not be offered by other sellers (for instance a fast recorded delivery). Unfortunately this information is not available in our dataset.

Each seller can also write a text description of the advert. We do not observe this text in our data. This additional piece of information could play a role if, for instance, for the same "as good as new" CD, one seller claimed that he only listened to it once and another seller claimed the CD was still unopened. These two potential sources of unobserved heterogeneity will be ignored in the benchmark analysis. We will however estimate an extension of the model that allows for unobserved heterogeneity in Section 7.

3.3 Descriptive statistics

As mentioned above, the estimation sample is taken from the third quarter of 2007. The reason why we restrict ourselves to a short time window is that the site has known a rapid growth rate and our estimation strategy assumes that beliefs regarding the joint distribution of characteristics and prices are constant.

We use data on sales of CDs with a catalog price ranging from €10 to €25 (hence leaving out EPs, CD singles or collector CDs). We discard transactions for which there was only one advert posted on the screen, as well as transactions for which one of the posted adverts had a price outside the range €1-20 (7% of adverts have a price outside this interval). This leaves us with 77,753 transactions, involving 23,538 sellers, 25,818 products and 145,823 adverts.[20]

The distribution of the number of adverts on the screen at the time of each transaction is presented in Table 1. We note that the majority of transactions occurred while there were less than 5 adverts available on the screen but that there are also many transactions for which the consumer had to choose from a large

[19] The construction of the dataset with all live adverts at each transaction date hinges on some assumptions which we present and discuss in Appendix C.

[20] Any statistics based on the population of adverts will be produced on a sample where each advert is counted only once, corresponding to the first time it appears in our sample.

number of adverts. Recall that the search cost and preference parameters are allowed to be heterogenous across transactions. We will show a breakdown of our estimation results by the number of adverts per transaction so that we can assess whether search frictions vary with the number of adverts.

Table 1: Distribution of the number of posted adverts per transaction

# adverts	Frequency	Percentage	Cumulated Percentage
2	19,248	24.76	24.76
3	13,234	17.02	41.78
4	9,597	12.34	54.12
5	6,975	8.97	63.09
6	5,415	6.96	70.05
7	4,106	5.28	75.33
8	3,244	4.17	79.51
9	2,645	3.40	82.91
$[10, 19]$	10,545	13.56	96.47
≥ 20	2,744	3.53	100.00
Any	77,753	100.00	100.00

A first striking observation from our sample is that, even though each advert or transaction refers to a specific CD (as defined by its barcode), we observe substantial price dispersion, in both the populations of adverts and of transactions. Table 2 shows that the average number of advert prices per product is above 5 (see last row). Comparing the first and the third columns, we also note that, for a given product, there are almost as many advert prices as there are adverts. Not only are advert prices different for a given product but they are also spread over a large support. The last column of Table 2 shows that on average the highest advert price is more than twice as large as the lowest one. This ratio increases substantially if we focus on products for which there are many adverts (for example, the ratio is above 3 if there are more than 6 adverts).

Table 2: Number of adverts, advert prices and advert price dispersion per product

# adverts per product	Frequency	Average number of advert prices	Mean $p_{\max}/p_{\min}$
2	7,015	1.95	1.49
3	4,760	2.89	1.92
4	3,267	3.80	2.30
5	2,384	4.69	2.71
6	1,724	5.57	3.03
7	1,270	6.47	3.35
8	947	7.29	3.56
9	784	8.09	3.77
$[10, 19]$	2,934	11.4	5.02
≥ 20	733	21.0	7.22
Any	25,818	5.07	2.69

Price dispersion is also substantial if we consider transactions. Table 3 focuses on products sold at least twice during our observation period. It shows that a given product, during a period of three months, can be sold on average at more than 3 different prices. Products with more than 8 transactions were on average sold at least at 5 different prices. Looking at the last column, we see that on average the highest transaction price is 76% higher than the lowest one for the same product. This relative difference rises above 100% for products sold more than 5 times.

Table 3: Number of transactions, transaction prices and price dispersion per product (sold at least twice)

# transactions per product	Frequency	Average number of transaction prices	Mean $p_{\max}/p_{\min}$
2	5,401	1.79	1.38
3	2,971	2.47	1.62
4	1,863	3.12	1.82
5	1,118	3.66	2.00
6	790	4.27	2.16
7	554	4.73	2.26
8	415	5.28	2.37
9	300	5.83	2.21
$[10, 19]$	891	7.60	2.67
≥ 20	244	14.40	2.86
Any	14,547	3.24	1.76

Table 4 shows how often the cheapest advert on the screen is chosen by consumers. We see on the last row that for 51.5% of all transactions, the advert that was sold was not the cheapest. Also, when consumers do not buy the cheapest advert, they choose an advert which is on average 56% more expensive than the cheapest advert. The other rows show that, as the number of adverts increases, the cheapest advert becomes less likely to be chosen and, when it is not chosen, its price is relatively lower than the price of the advert actually bought by the consumer.

Table 4: Cheapest and sold advert prices per transaction

			If $p_{\text{sold}} > p_{\min}$	
# adverts	Freq.	% $\{p_{\text{sold}} = p_{\min}\}$	mean $\frac{p_{\text{sold}}}{p_{\min}}$	rank p_{sold}
2	19,248	67.96	1.35	2.00
3	13,234	55.26	1.41	2.29
4	9,597	47.75	1.47	2.54
5	6,975	42.09	1.52	2.74
6	5,415	39.48	1.53	2.92
7	4,106	37.46	1.56	3.06
8	3,244	36.65	1.61	3.29
9	2,645	35.54	1.62	3.37
$[10, 19]$	10,545	31.76	1.80	3.84
≥ 20	2,744	23.87	2.05	5.02
Any	77,753	48.51	1.56	2.94

These stylized facts on price dispersion may have three different explanations: seller/advert characteristics, consumer preferences for advert characteristics and search costs. Each of these dimensions can exhibit some heterogeneity. In our model, they are captured by x, γ and s respectively.

In Appendix E we show that, in addition to price dispersion, there is also variation in the non-price characteristics (reputation, size, condition and seller status) among adverts and among transactions, even if we control for price. If consumers care for these characteristics, this source of differentiation could explain some of the price dispersion even in a context of perfect information. We now show however that search frictions also play a role on this Internet platform, and that heterogeneity in advert characteristics and in consumer preferences cannot fully explain the dispersion in prices.

We saw above that in around 52% of transactions, the cheapest advert was not the one chosen by the consumer (see Table 4). The consumer's choice may thus be driven by non-price advert characteristics and/or hindered by search frictions. To motivate the inclusion of search frictions in our analysis, we compute the fraction of transactions which took place in the presence of an 'unambiguously better' advert. We say that an advert j is 'unambiguously better' than i if j is at least 'as good as' i in terms of price, seller reputation and product condition and strictly better in at least one of these dimensions and if the sellers of adverts i and j have the same status (professional or individual) and are in the same size category.[21] This definition does not depend on any assumptions on consumers' preferences for seller size or status and only assumes that consumers' utility weakly increases with lower prices or better reputation or better product condition.

Counting transactions where an alternative advert was 'unambiguously better' than the advert bought gives a very conservative lower bound on the number of transactions that cannot be explained without search costs, even when heterogenous preferences for observed advert characteristics are allowed for. The proportion of transactions with 'unambiguously better' adverts is shown in Table 5. We see that this proportion is higher than 7% on average and that it increases with the number of adverts on the screen. This motivates the introduction of consumer search costs.[22]

Table 5: Transactions where an 'unambiguously better' advert was available

# adverts	Freq.	% 'unambigously better'
2	19,248	4.28
3	13,234	5.83
4	9,597	6.88
5	6,975	8.00
6	5,415	8.57
7	4,106	8.13
8	3,244	9.86
9	2,645	8.96
$[10, 19]$	10,545	10.21
≥ 20	2,744	14.07
Any	77,753	7.24

4 Empirical strategy

Our empirical approach mainly consists of three steps. First we project the vector of advert characteristics onto a scalar variable X, referred to as the hedonic index. Secondly, we estimate the consumers' beliefs about the distribution of X given price. Lastly, we compute the sets of parameters (s, γ) that can rationalise each transaction. In this section, we present each of these three steps.

[21] Where we define four size categories as: $[0, 50[$, $[50, 500[$, $[500, 5000[$ and ≥ 5000.

[22] Another explanation for the statistics in Table 5 would be the presence of unobserved advert characteristics, a case we will consider at the end of the paper in section 7.

4.1 Aggregation of advert characteristics

As seen in Section 2.4, handling the consumer problem with only two dimensions of choice x and p greatly improves the tractability of the search problem. With a scalar x, the characterisation (9)-(10) of the identified sets consists of a simple comparison of instantaneous and reservation utilities, the latter being easily obtained, through expression (8). We think that the gains in clarity and tractability of working with a scalar hedonic index outweigh its only drawback, which is the lack of heterogeneity in the marginal rate of substitution between two non-price characteristics (recall that we still allow for heterogeneity in the MWP for the scalar hedonic index).

Advert characteristics and the hedonic index. We define our aggregate hedonic index x as a linear projection of the advert characteristics. The index x is thus a weighted sum of the different advert characteristics and we denote as β the vector of weights, which is constant across consumers. Preference heterogeneity is now reduced to one dimension, embodied by the scalar parameter γ, the MWP for x, which is heterogeneous across consumers.

The advert characteristics to be projected onto x are the following: five seller size dummies (≤ 50, $[51, 100]$, $[101, 500]$, $[501, 5000]$, > 5000), four product condition dummies ('good, 'very good, 'as new', 'new'), a professional seller dummy and the seller's reputation.[23]

We set the β coefficient of the reputation variable to be equal to 1 (if all projection weights were free parameters, we could not identify β and γ separately). We also make the very weak assumption that consumers have non-negative preferences for seller reputation. Note that this ensures that $\gamma \geq 0$ and that γ can then be interpreted as the marginal willingness to pay for a 0.1 increase in reputation.

The distribution of x given price, and thus the ψ function defined in section 2.4, depend on the value of β. Hence, the first two steps of our empirical approach are interrelated. For the moment, let us assume that we can compute (an estimate of) ψ for any value of β so that we can focus on the estimation of this parameter. The next section will then show how we estimate ψ for a given β.

Estimation of the hedonic index. In this section only, we focus on transactions where $J = 2$ adverts are available. This will substantially ease the computational burden of estimating β and the value we will find can be applied to other transactions since β is assumed to be homogenous across consumers.

For a given parameter value β (and the corresponding belief function) we can define the dummy indicator $pass_i(\beta)$ which is equal to 1 if there is at least one value of (s, γ) that can rationalise transaction i and 0 otherwise. In other words, and using the results from sections 2.3-2.4, $pass_i(\beta)$ equals 1 if and only if there is at least one value of (s, γ) such that (3) holds (if $s = 0$), or (4) holds (if $\gamma = 0$), or neither (9) nor (10) holds (if s and γ are both strictly positive).

Focusing on transactions with two adverts makes the computation of $pass(\beta)$ manageable. Indeed, if $p_i \leq p_j$ then the transaction is explained by $\gamma = 0$ (the consumer only cares about prices). Recall that we

[23] We use a variable equals to 10 times the seller's reputation. Most of the variation in reputation is between 4 and 5. The very few sellers with reputation levels below 4 or with no reputation yet (no completed transactions) have been assigned a reputation equal to 4 (thus a reputation variable value of 40).

use index i for the advert sold and index j for an alternative advert (available but not sold). If $p_i > p_j$ and $\overline{\gamma}\,(x_i - x_j) > p_i - p_j$ then it is explained by $s = 0$ (advert i offers more utility than j). We denote by $\overline{\gamma}$ the highest value for the MWP that we consider intuitively plausible. In our aplication, we set this to 20, which means that the maximum price difference that consumers are prepared to pay for a 0.1-increase in reputation is 20 euros. As this corresponds to the highest price of a CD in our sample, it seems to be an uncontroversial upper bound for the value of $\overline{\gamma}$.[24] If we cannot explain the transaction by either $s = 0$ or $\gamma = 0$ then advert j yields a strictly higher utility than advert i so the parameters must be such that the consumer drew i first and then did not look at j (because of search frictions). Formally this means that $r_i \geq r_j$ and $u_i \geq r_j$. For a given value of β, we then browse a grid of values for s and γ and set $pass_i(\beta)$ to one when both these inequalities are verified for at least one value of (s, γ) (more details on this grid are given in section 4.3).

If $pass_i(\beta) = 1$, we can define $\underline{s}_i(\beta)$ as the lowest search cost that explains transaction i and take its average across explained transactions: $M(\beta) = E\left[\underline{s}_i(\beta)|pass_i(\beta) = 1\right]$. We then estimate β by maximising the following criterion:

$$C(\beta) = \sum_{i,J=2} pass_i(\beta) + e^{-M(\beta)}. \tag{11}$$

With this criterion, our estimate of β achieves two objectives. First, it maximises the number of transactions with two adverts that can be explained by our model (thus the first term on the right-hand side of equation (11)). This is not enough to point-identify β however as the argmax may not be a singleton. We thus add a second term to $C(\beta)$ which minimises the average minimal search cost required to explain the transactions. Note that the first term is an integer (sum of dummy variables) whilst the second term is between 0 and 1. Hence our criterion allows us to minimise the average minimal search cost without departing from the maximum number of transactions fitted by the model. Also, note that if $p_i \leq p_j$, then $pass_i(\beta) = 1$ for any β since $\gamma = 0$ can rationalise the transaction. These transactions are hence not useful in our identification of β.

Our motivation for this method of estimating β is twofold. Our first priority is to maximise the fit of the model. This, however, does not pin point a unique estimate of β. We then proceed to point-identify it by minimising the necessary search costs to account for transactions with only two adverts. The spirit of this choice is to give a fair chance to the perfect-information model to fit our data and hence to interpret our findings regarding search costs as credible evidence of the minimum amount of search frictions needed to explain the data.

4.2 Estimation of consumers' beliefs

As seen in section 2, consumers' beliefs regarding the joint distribution of (X, P) is central to their search and purchasing behaviour. We now explain how we compute these beliefs for a given projection β. Beliefs could either be characterised by the conditional distribution of X given price, $F_{X|P}$, by the ψ_p functions (one for each price) or by the inverse of these functions, ψ_p^{-1}. The latter quantity is the main target as it is

[24] An alternative would be to not impose an upper bound for γ and say that the transaction is explained by $s = 0$ as soon as $x_i > x_j$.

the one we use to compute reservation utilities (cf. equation (8)).

We consider that we are in an equilibrium where consumers' beliefs about the joint distribution of P and X coincide with the actual distribution of these two variables in the population of adverts. This distribution is observed in the data and can thus be estimated non-parametrically. In what follows, we assume that a consumer's beliefs about X depend on prices only up to the first decimal.[25] In the following estimation of beliefs, the data is grouped by price rounded to the nearest first decimal.

Since all but one of the advert characteristics (reputation) are dummy variables, and reputation essentially takes one out of eleven values (from 4.0 to 5.0), the support of X, which we denote as $\mathcal{X}$, can be seen as discrete. Given the large number of adverts in our data, we compute a non-parametric estimate of the ψ_p functions as follows:

$$\widetilde{\psi}_p(x) = \sum_{\widetilde{x} \in \mathcal{X}, \widetilde{x} > x} (\widetilde{x} - x) \cdot \widehat{f}\left(\widetilde{x} | p\right), \tag{12}$$

where $\widehat{f}\left(\widetilde{x}|p\right)$ is the number of adverts at price p with $X = \widetilde{x}$ divided by the number of adverts at price p.

We can then invert this function[26] and obtain a "raw" estimate of ψ_p^{-1} used in the computation of the reservation utilities. However this approach leads to ψ functions which are quite jagged. For instance, the dependence of $\psi_p(x)$ on p for a given x is clearly increasing overall but it shows fluctuations for some prices probably driven by a few adverts in small samples for a given price.

We thus proceed by smoothing these "raw" beliefs and fit the "raw" function $(p, y) \longmapsto \widehat{\psi}_p^{-1}(y)$ with a 3rd-order polynomial and impose the derivative with respect to y to be negative (by definition of the ψ function) and the derivative with respect to p to be positive. The latter restriction means that consumers expect the distribution of X to improve when the advert price increases.[27] The resulting "smooth" estimate is denoted as $\widehat{\psi}_p^{-1}$.[28] We will use the "smooth" beliefs as our benchmark but for completeness we also present, in Appendix F, estimation results using the "raw" beliefs.

4.3 Grid search of preference and search cost parameters

The last stage of our estimation procedure is relatively straightforward. We use the scalar hedonic index and the $\widehat{\psi}_p^{-1}$ functions estimated in sections 4.1-4.2 to compute the utility u and reservation utility r for each advert. In the spirit of the revealed-preference approach, we then browse a two-dimensional grid for s and γ. For s the grid values go from 0 to 5 with a step of 0.05, then from 5 to 10 with a step of 0.1.[29] The grid values for γ are the same as for s except that we also include the values 15 and 20. At each grid point, we check whether the parameter values (s, γ) are consistent with each transaction i using conditions (3)-(4) and (9)-(10). The first condition is used when $s = 0$, the second when $\gamma = 0$ and the last two if s and γ are both strictly positive.

[25] This is done for practical reasons and we think it is not unrealistic to assume that buyers may expect the same characteristics from two adverts with prices at €10.64 and €10.61.

[26] Which is monotonic by construction.

[27] As explained in section 2.4, by "improve" we mean that $F_{X|P=p'}(\cdot)$ 2nd-order stochastically dominates $F_{X|P=p}(\cdot)$ when $p' > p$.

[28] Note that since ψ_p^{-1} is approximated by a polynomial, the constraints we impose on the derivatives are linear in the coefficients thus making the constrained optimisation problem easily tractable.

[29] Our choice of upper bound for the values of s, i.e. 10 euros, seems reasonable as it represents the maximum disutility a consumer can experience from sampling an additional advert. 10 euros is also half of the maximum price of CDs in our sample, and thus is unlikely to understate the maximum search cost on this market.

We have decided to resort to a grid-search because the inequalities defining the identified set for each transaction are non-linear in the parameters of interest. An issue of this approach is that we could miss pass values between two grid points. We try to avoid this issue by using a relatively fine grid and relying on the smoothness of the functions entering inequalities (3)-(4) and (9)-(10) – indeed recall that we use the "smooth" specification of beliefs in our benchmark estimation.

Our empirical approach will produce sets of possible values for the search cost and preferences parameters for each transaction. Our parameters are not point-identified within these sets so there is modelling uncertainty. However these sets exhibit no sampling variability, and this for two reasons. First, we are observing the whole population of interest in our data (consumers who purchase a CD on the PriceMinister website during the observation period). Second, the inequalities (3)-(4) and (9)-(10) only involve observed variables and are deterministic.[30] Uncertainty on the resulting sets, which would call for computing standard errors, would arise if we wished to draw inference on a larger population of consumers (of which our exhaustive data of transactions on PriceMinister would be a subsample) or if we allowed the utility function to depend on an unobserved stochastic taste shock.

5 Estimation results

We start this section with a look at the fit of the model and then show our results on the preference and search cost parameters. We also show how these estimated sets of parameters can be used to infer on heterogeneity in the population of buyers. These results will be based on the benchmark specification of consumers' beliefs (i.e. the "smooth estimation" of beliefs as explained in section 4.2). Results for the "raw" specification of beliefs are presented in Appendix F.

5.1 Model fit

For each transaction, we will consider that our model fits the observed choice if S_i is not empty i.e. there is at least one value of (s, γ) such that we cannot reject the model with conditions (3)-(4) or (9)-(10) for all the (i, j) comparisons relevant to this transaction.

Table 6 shows the fit of our model. In this table, and in many tables in subsequent sections, results will be broken down by the number of adverts per transaction. We also group transactions into two categories, depending on whether an alternative 'better' advert was available but not bought. Recall that an advert j (not sold) is 'better' than advert i (sold) if it is both cheaper and offering a higher hedonic index, i.e. $p_j \leq p_i$, $x_j \geq x_i$ with at least one of these two inequalities being strict. These transactions with 'better' adverts play an important role in the identification of positive search costs. The last row of Table 6 shows that 16,710 out of 77,753 (21.49%) of all transactions had at least one 'better' advert.

The main result from Table 6 is that our model explains almost all transactions (95.36%), whether the number of adverts was small (99.21% if 2 adverts) or large (90.69% if more than 15 adverts). The fit is even higher (99.8%) among transactions with no 'better' advert. The fit is still high (79.16%) if we look

[30] The first step of our estimation does not use the whole population but only transactions with two adverts. There may a sampling design issue there if the projection parameter β is not homogeneous across consumers. However challenging this assumption is difficult in practice as this first step is computationally demanding.

at transactions with at least one 'better' advert and remains relatively stable when the number of adverts per transaction increases. Our search model thus does a very good job at explaining the transactions on this Internet platform, even when considering transactions with 'better' adverts, that is transactions that would not be explained by a perfect information model (unless the consumer only cares about prices and the 'better' and sold adverts have the same price).[31]

Table 6: Pass rate (%) among transactions with J adverts

	All transactions		Transactions with no 'better' advert		Transactions with $\geq$ 1 'better' advert	
# adverts	Freq.	% Pass	Freq.	% Pass	Freq.	% Pass
2	19,248	99.21	17,180	100.00	2,068	92.65
3	13,234	97.23	11,100	99.91	2,134	83.27
4	9,597	95.69	7,579	99.72	2,018	80.53
5	6,975	94.74	5,309	99.64	1,666	79.11
6	5,415	93.63	3,975	99.72	1,440	76.81
7	4,106	93.98	3,045	99.80	1,061	77.29
8	3,244	92.48	2,308	99.74	936	74.57
9	2,645	93.31	1,866	99.57	779	78.31
10	2,174	92.00	1,515	99.54	659	74.66
11	1,707	91.10	1,173	99.49	534	72.66
12	1,472	91.51	999	99.50	473	74.63
13	1,192	89.35	794	99.62	398	68.84
14	1,063	90.69	698	99.71	365	73.43
$\geq$ 15	5,681	90.69	3,502	99.49	2,179	72.60
Any	77,753	95.36	61,043	99.80	16,710	79.16

Freq: number of transactions in each category.

% Pass: proportion (in %) of transactions that can be explained - $S_i \neq \{\emptyset\}$.

5.2 Search costs

We now quantify the role of search costs on the platform. We have just shown that our model rationalises almost all transactions in our sample. We now show that many of these transactions could not be explained without strictly positive search costs and produce bounds on the distribution of search costs among explained transactions.

Strictly positive search costs. For each transaction i that is explained by the model, we can define $\underline{s}_i$ as the lowest value of s that allows our model to be consistent with the observed transaction i. We can then look at the average of $\mathbf{1}\{\underline{s}_i > 0\}$ –an indicator that a strictly positive search cost is needed to account for transaction i– among all transactions explained by the model. Results are shown in Table 7.

We see that strictly positive search costs are needed for 22.34% of all transactions explained by our model. This proportion increases steeply with the number of adverts per transactions (9.37% if there are 2 adverts, 43.74% if there are more than 15 adverts). This is expected as when the number of adverts is large, it is more likely that a 'better' advert will be available which then requires a positive search cost for the model to explain the transaction. An interesting result which does not appear in Table 7 is that the minimum search

[31]We discuss in Appendix G the power of our approach, i.e. its ability to reject the model with a dataset where choices are made at random.

cost conditionnally on transaction i being explained and $\underline{s}_i$ being strictly positive does not vary much with the number of adverts. In other words, strictly positive search costs are more often needed to fit the data when there are more adverts on the screen but the magnitude of these costs does not increase.[32]

For transactions with no 'better' adverts, positive search costs are rarely needed (6.66%) unless the number of adverts is large. This is not surprising as most transactions with no 'better' advert can be explained by a perfect-information model and the appropriate taste for x. This is not systematic though, as we discussed in Section 2.4, and our results suggest that there is a non-negligible proportion of transactions for which the MWP's implied by comparing the sold advert with different alternatives are not consistent, so that the transaction can only be explained with a positive search cost.

Table 7: Proportion (%) of explained transactions requiring strictly positive search costs

	All transactions		Transactions with no 'better' advert		Transactions with ≥ 1 'better' advert	
# adverts	Freq.	% $\underline{s} > 0$	Freq.	% $\underline{s} > 0$	Freq.	% $\underline{s} > 0$
2	19,096	9.37	17,180	0.15	1,916	92.12
3	12,867	15.12	11,090	2.54	1,777	93.64
4	9,183	21.17	7,558	5.45	1,625	94.28
5	6,608	25.61	5,290	8.54	1,318	94.08
6	5,070	28.46	3,964	9.71	1,106	95.66
7	3,859	29.67	3,039	11.42	820	97.32
8	3,000	32.57	2,302	13.21	698	96.42
9	2,468	35.49	1,858	16.15	610	94.43
10	2,000	34.55	1,508	14.46	492	96.14
11	1,555	36.66	1,167	17.48	388	94.33
12	1,347	38.38	994	18.41	353	94.62
13	1,065	36.81	791	16.44	274	95.62
14	964	37.97	696	16.38	268	94.03
≥ 15	5,066	43.74	3,484	20.09	1,582	95.83
Any	74,148	22.34	60,921	6.66	13,227	94.57

Freq: number of transactions in each category.
%$\underline{s} > 0$*: proportion (in %) of explained transactions that need a strictly positive search cost.*

The overwhelming majority of explained transactions with a 'better' advert require a strictly positive search cost (94.57%). Not all of them do because some of the 'better' transactions are such that $p_i = p_j$ and $x_i < x_j$: these can be explained without search costs by relying on a zero taste for the hedonic index ($\gamma = 0$) and our rule for ties. In any other case, a transaction with a 'better' advert cannot be explained unless $s > 0$.

We view the results from Table 7 as particularly important as they provide strong empirical evidence of strictly positive search costs on this Internet platform and yet are based on a very flexible specification of consumers' preferences (we allow for full heterogeneity in γ and impose no parametric assumption on its distribution) and of their search behaviour, where the sampling order is shaped by preferences and search costs.

[32] The average is 2.75 for explained transactions with $\underline{s} > 0$ and any number of adverts, and this average is 2.4 (resp. 2.7, resp. 2.2) when there are 2 (resp. 10, resp. more than 15) adverts.

Bounds on the distribution of search costs. We now infer on the distribution of s in the population of transactions that can be explained by our model (95.36% of our sample). We denote the set of transactions i that can be fitted by our model as I_f and the number of these transactions as N_f. Let $\underline{s}_i$ be the minimum search cost that makes the model consistent with a given transaction i. We can also define $\overline{s}_i$ as the highest search cost that fits this transaction. The lower bounds exist as S_i is not empty (we only consider explained transactions). The upper bounds are limited by the highest value on our grid.

For each value of s, we define the lower and upper bounds of the c.d.f. of s across explained transactions as:

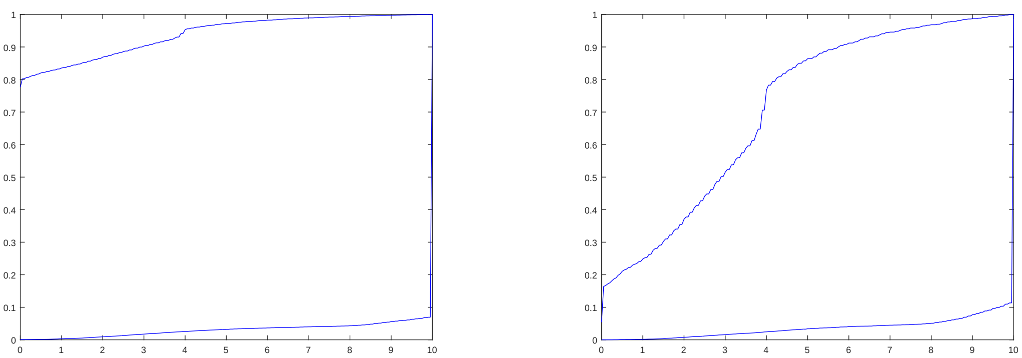
$$\frac{1}{N_f} \cdot \sum_{i \in I_f} \mathbf{1}\left\{\overline{s}_i < s\right\} \quad \text{and} \quad \frac{1}{N_f} \cdot \sum_{i \in I_f} \mathbf{1}\left\{\underline{s}_i < s\right\} \qquad \text{respectively.} \tag{13}$$

We also compute these expressions among explained transactions with at least one 'better' advert and obtain bounds on the c.d.f. of s for these specific transactions.

The bounds on the c.d.f. of search costs are shown in Figure 1, in the population of all explained transactions and of explained transactions with 'better' adverts. We first note that in both cases the lower bound is close to (but different from) 0. This is not systematic though as we can see that the lower bound is not flat. The second remark is that if we look at all explained transactions, the bounds we find are wide. This was expected as we know from Table 7 that 1-22.34%=77.66% of transactions can be explained with search costs equal to 0. However, if we focus on transactions with 'better' adverts, we get tighter bounds and search costs can be relatively high, with a median at 2.9 for $\underline{s}$. Hence minimum search costs over the whole population of transactions are low (they are zero for 78% of transactions) but, for those who have a positive search cost, the minimum cost is high relative to the average price of items traded. Note that one could draw a parallel between this result and the equilibrium consumer search models which, following Stahl II (1989), assume that there is a mass point of consumers with zero search cost ("shoppers") and other consumers with strictly positive search costs.

Figure 1: Bounds on the c.d.f. of the search cost s

All explained transactions | *Explained transactions with ≥ 1 better advert*



Vertical axis: lower and upper bounds - Horizontal axis: $\underline{s}$

5.3 Preferences

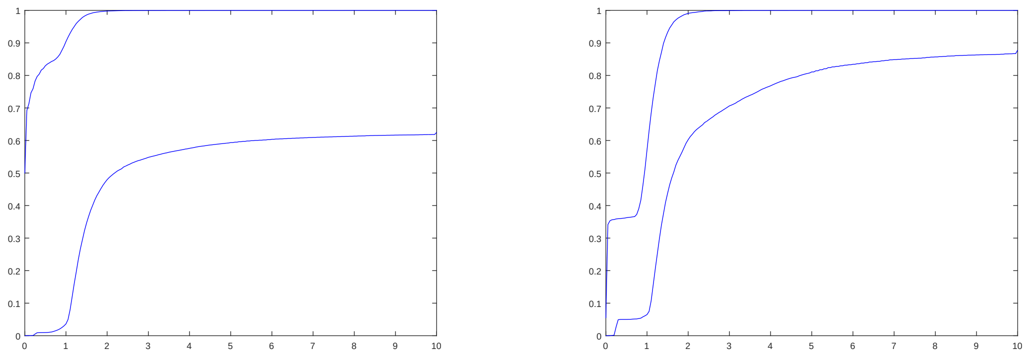
The estimated scalar hedonic index. Table 8 shows the estimated value of the β parameter i.e. the projection of advert characteristics onto a scalar index x (see section 4.1). We note that the item condition dummies have signs and relative magnitudes that are consistent with intuition (the omitted category is 'new'). The resulting aggregate hedonic index x ranges from 20.59 and 50.51, and its distribution in the population of adverts has an average of 36.23, a standard error of 6.67 and a median at 35.84. With this scalar index, a consumer with a MWP $\gamma = 1$ would be willing to pay about €4 to go from the 25% percentile (31.78) to the median (35.84) of the advert quality distribution. As a comparison, recall that $\gamma = 1$ corresponds to a marginal willingness to pay €1 for a 0.1 increase in seller's reputation.

Table 8: Estimation of the β parameter

Reputation (normalized)	1.000
Size $\in [50, 100]$	-0.650
Size $\in [101, 500]$	0.186
Size $\in [501, 5000]$	1.021
Size > 5000	2.006
as new	-11.178
very good	-16.221
good	-18.757
Professional	-0.508

Bounds on the distribution of preference parameters. We already know from Table 4 that only less than half of the transactions could be explained by a model with $\gamma = 0$ (since the cheapest advert is bought in 48.5% of transactions). We compute bounds on the c.d.f. of γ in the population of explained transactions using the same method as the one used for search costs in section 5.2 (see equation 13). These are shown on Figure 2. These bounds on the distribution of γ are relatively tight. The median of the MWP for x is between 0 and 2.2 and, if we focus on transactions with 'better' adverts, between 0.9 and 1.6. Given that we are considering CDs with an advert price between €1 and €20, this is evidence that consumers are willing to pay substantially for non-price advert characteristics and that this consumers' taste for the hedonic index of quality plays an important role in explaining the transactions observed in our data. This in turn is suggestive of an environment where search costs may play a role since non-price characteristics are less readily observable than prices. Our modelling choice to assume that non-price characteristics are costly to observe is key in this complementarity between our findings on preferences and our findings on search costs.

Figure 2: Bounds on the c.d.f. of the preference parameter γ



Vertical axis: lower and upper bounds - Horizontal axis: $\underline{\gamma}$

5.4 Heterogeneity in preferences and search costs

We now show that, even though parameters s and γ are set identified for each transaction, we can use our results to provide evidence on the degree of heterogeneity in preferences and search costs. Our model explains N_f transactions. For each of these we obtain a set S_i that is non-empty. Within these sets, our accept/reject approach is agnostic as to where the individual parameters (s, γ) lie. These parameters are set identified for each individual transaction.[33] If we assumed homogeneity, we would then be able to aggregate all sets S_i to set identify the preference parameters (γ, s) for all individuals as the intersection of all S_i across all explained transactions. However, we allow for heterogeneous preferences so that we have set identification of (γ, s) for each explained transaction and we cannot take the intersection of these sets to identify our parameters. Still, we can recover some information on the underlying heterogeneity distribution that is compatible with this collection of individual sets S_i.

To this end we borrow from Stoye (2010) and compute bounds on the variance of the distribution of γ (respectively s) compatible with the aggregation of the sets S_i over the N_f transactions that our model is able to rationalise. This provides bounds on the extent of unobserved heterogeneity in γ (respectively s) in our data. Admittedly, the bounds that we obtain on the variance of the underlying heterogeneity distribution can be wide. Nevertheless we think that our estimation of the minimum amount of heterogeneity in s and γ compatible with our dataset is informative.

The procedure is the following. Consider the lower and upper bounds on the c.d.f. of γ, respectively denoted as $\underline{H}$ and $\overline{H}$, that we computed and illustrated in Figure 2. Following Stoye (2010) we know that the most compressed distribution compatible with these bounds will be a distribution of type H_c such that:

$$H_c(\gamma) = \underline{H}(\gamma) \cdot \mathbf{1}\{\gamma < c\} + \overline{H}(\gamma) \cdot \mathbf{1}\{\gamma > c\},$$

[33]We cannot ensure that each consumer buys only one item during our observation period. In this section however, we will consider that heterogeneity relates to individuals, thereby ignoring the possibility that several transactions may correspond to the same individual.

and that the most dispersed distribution compatible with the identified bounds will be of type H_d such that:

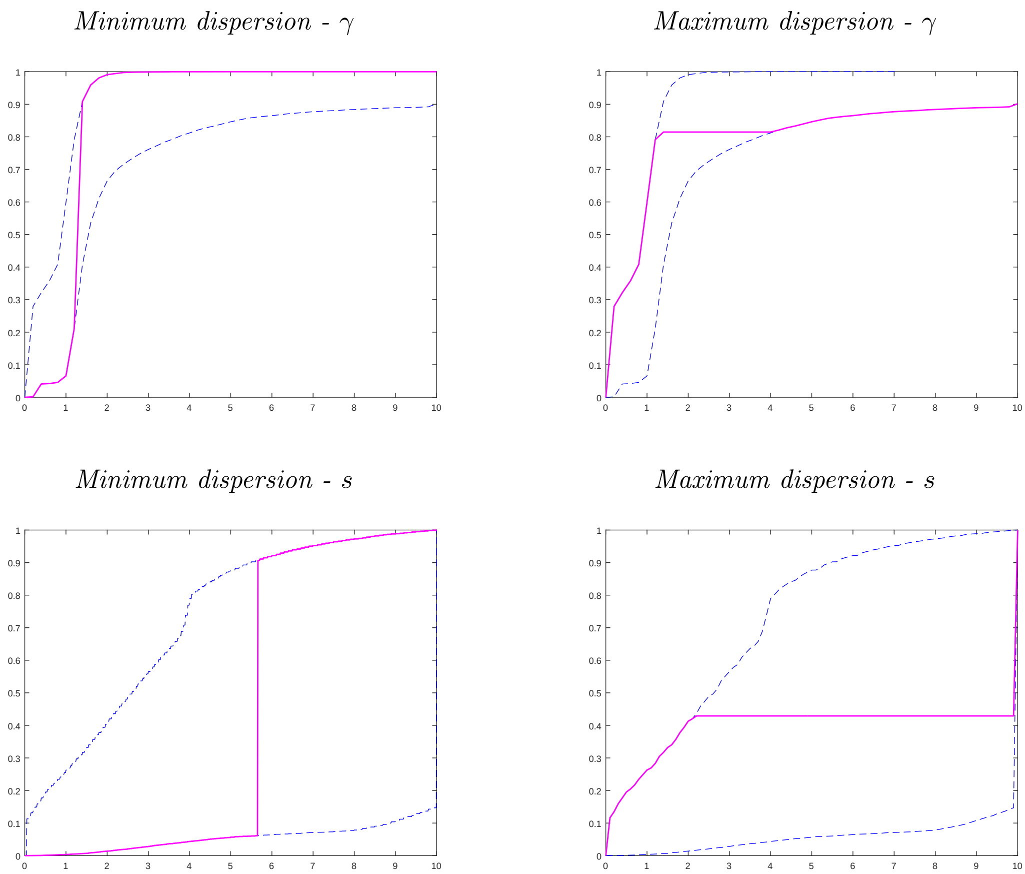
$$H_d(\gamma) = \overline{H}(\gamma) \cdot \mathbf{1}\left\{\overline{H}(\gamma) < d\right\} + d \cdot \mathbf{1}\left\{\underline{H}(\gamma) < d < \overline{H}(\gamma)\right\} + \underline{H}(\gamma) \cdot \mathbf{1}\left\{\underline{H}(\gamma) > d\right\}.$$

Based on our estimates, these distributions are shown on Figure 3 for values that minimise the variance of H_c with respect to c and maximise the variance of H_d with respect to d. These values, respectively denoted as c^* and d^*, are characterised by:[34]

$$c^* = \int_{\gamma_{\min}}^{\gamma_{\max}} \gamma dH_c(\gamma) \quad \text{and} \quad \frac{1}{2}\left[\overline{H}^{-1}(d^*) + \underline{H}^{-1}(d^*)\right] = \int_{\gamma_{\min}}^{\gamma_{\max}} \gamma dH_d(\gamma). \tag{14}$$

We find that the standard deviation of the distribution of γ compatible with the bounds we obtained is between 0.14 and 9.4. The mean of the most compressed distribution of γ is 1.26. With this wide range and rather small lower bound for the standard deviation, we can rule out the homogenous case but we cannot ensure that there is a large amount of heterogeneity in consumers' MWP since the minimum standard deviation is small (indeed it is equal to 0.14 thus nine times lower than the mean to 1.26). Note that our result on the maximum standard deviation is less robust since it is sensitive to our choice of upper bound for the range of γ. This is not crucial for our analysis however as detecting a positive amount of consumer heterogeneity hinges on the lower bound of the standard deviation.

Figure 3: Bounds on heterogeneity of s and γ



Dashed: lower/upper bound on cdf - Solid: Min/Max dispersion cdf

[34] Details of this derivation are available upon request.

We can use the same procedure to quantify the heterogeneity in search costs s. The standard deviation of the distribution of search costs s is bounded between 0.58 and 4.95, which means that the minimum amount of heterogeneity with respect to search costs compatible with our data is more substantial. Accordingly, an eye-ball comparison of the two graphs in the left column of Figure 3 shows that the minimum-variance distribution of search costs shows thicker tails than its counterpart for the preference parameter. We think this is an interesting feature of our results, both as a descriptive tool and as an input to future work on general equilibrium models of Internet trade. Indeed, this dispersion of search costs is relevant from the sellers' point of view, who can devise their pricing decision in the presence of potential buyers with different search costs.

6 Search strategies

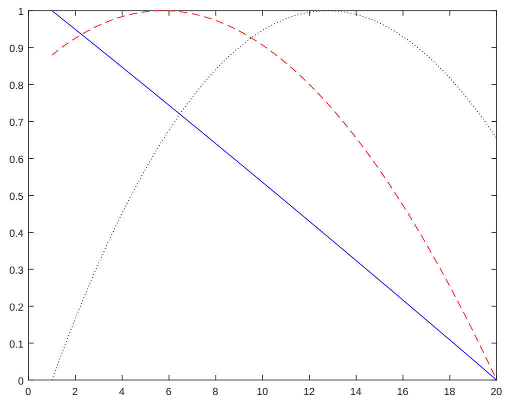
In this section we look at consumer search patterns and illustrate how the structure of our model and the flexible modelling of heterogeneity in the (s, γ) parameters allows us to capture a rich set of search behaviours. This is a key strength of our modeling approach, which rationalises heterogeneity in consumers' sampling behaviour in a framework with directed search.

Sampling order. Using the β and ψ estimates from the first two steps of our benchmark estimation (see sections 4.1-4.2), we can compute reservation utilities r for any price p and any value of (s, γ). We can then study how the sampling order depends on consumers' preferences (γ) and search cost (s).

In Figure 4 we plot price against sampling order for different values of s and γ. Note that the following pertains to sampling behaviour rather than to purchases. As detailed in section 2, the optimal search strategy is to sample adverts in descending order of reservation utility. Our theoretical framework delivers rich predictions in terms of sampling order, of which we now give several illustrations. For each (s, γ), we compute $r(p, s, \gamma)$ for all prices on a grid from 1 to €20 (with a step of 0.1). The results of this computation show that consumers' sampling order is not necessarily monotonic in prices. Figure 4 displays the reservation value $r(p, s, \gamma)$ on the vertical axis versus the price on the horizontal axis. The reservation value has been normalised to be between 0 and 1 for Figure 4 only, in order to facilitate comparisons across different values of (s, γ) since our focus in this section is the sampling order. When, for example, the consumer samples items in increasing order of price, the sampling order will be represented by a decreasing line since r decreases with price. Non-monotonic sampling behaviour arises in some cases: the consumer starts sampling a mid-range price level where $r(p)$ reaches its peak and subsequently samples items higher and lower than the initial price in an alternating pattern predicted by the series of reservation utilities for the consumer's individual (s, γ). In this case, the figure shows sampled prices getting further and further away from the initial price level as the sampling process proceeds.

Figure 4: Price and search order, the effect of preferences and search cost

Horizontal axis: p - Vertical axis: $r(p, s, \gamma)$ (normalised)



Solid: $\gamma = 0.05$, $s = 0.05$ (or $s = 1$) - Dashed: $\gamma = s = 1$ - Dotted: $\gamma = 2$, $s = 1$

The solid line in Figure 4 shows the sampling order when the MWP for x is very low (€0.05) and the search cost equals €0.05 or €1. In either of these cases, when consumers barely care about the hedonic index x they sample adverts by increasing order of price. This is important for our results: the observed consumer behaviour needs to be explained by the *joint* presence of search costs and a consumer's taste for non-price characteristics.

The dashed line shows that, in the presence of search costs ($s = 1$), the search pattern markedly changes when preferences for x are stronger ($\gamma = 1$). The sampling order is no longer increasing in price i.e. consumers do not sample the cheapest advert first and carry on sampling in ascending price order. Indeed, we see that they would first look at adverts with a price around €6 then at prices of €5 or €7. The most expensive adverts are still sampled last.

Increasing further the MWP for x yields another search pattern. Looking at the dotted line, we note that, for a given search cost, $s = 1$, as the taste for the hedonic index γ increases to 2, the first price sampled by the consumer rises from €6 to €13. Another difference is that the most expensive adverts are no longer sampled last. This is intuitive as the more consumers care about hedonic advert characteristics, the more likely they are to look at expensive adverts before cheap adverts.

To summarise, Figure 4 illustrates the flexibility of our search model with respect to sampling patterns. Depending on parameter values, consumers may not sample adverts by increasing order of price. The sampling order and thus the consumer's choice set will depend on his preferences and search cost. As far as we know, ours is the first empirical analysis of this type of consumer behaviour in the context of a directed search model.

Sampling strategies. Using the identified sets of parameters (s, γ), we can delve further into sampling patterns. Since (s, γ) are only set-identified, for the purpose of our illustration, we need to select for each explained transaction i a value of (s, γ) in S_i. We choose to focus on the lowest search cost needed to explain the transaction, $s = \underline{s}_i$, and its corresponding lowest MWP, $\gamma = \underline{\gamma}(\underline{s})$.

For these values of (s, γ), we can compute for each transaction the utility u and reservation utility r of each advert. We can then see how often the sold advert i has the highest reservation utility, in which case it will be sampled first (or sampled first with positive probability if other adverts have the same reservation utility). We also look at the sampling order of the advert sold i, which will be given by the rank of r_i –in other words the number of adverts sampled before the chosen one. We report the average and maximum sampling order. Lastly, we check whether the consumer keeps sampling adverts once he has drawn the advert that he will eventually buy. According to the model shown in Section 2, this will happen if there is an alternative advert j such that $u_j \leq u_i < r_j < r_i$. For these statistics to make sense, we focus on the 16,565 transactions that can be explained by the model and for which search costs are strictly positive, $\underline{s} > 0$.

Results are shown in Table 9. We see that the sold advert has the highest reservation value (and is thus sampled first) 63% of the time on average. This figure decreases sharply with the number of adverts. The next column shows that on average, the sold advert is sampled early. This is not systematic however, as shown by the next column. Indeed, adverts sold can have a high sampling order, as the maximum rank of r_i is often close to the actual number of adverts on the screen. This means that in some cases, the sold advert was sampled after most of the other adverts. The last column of Table 9 shows that consumers may carry on sampling adverts after they have drawn the one that they will eventually buy. This is consistent with the search patterns observed by De Los Santos et al. (2012) albeit in a different context (search for books across e-commerce websites). As discussed in Section 2 and now shown with this set of results, such a consumer behaviour is consistent with our sequential directed search model.

Table 9: Sampling strategies, explained transactions with $s = \underline{s} > 0$

# Adverts	Freq.	% Samp. 1st	Mean	Max	% Keep Samp.
2	1,790	99.55	1.00	2	2.85
3	1,946	78.62	1.25	3	4.21
4	1,944	67.70	1.42	4	8.23
5	1,692	63.53	1.57	5	11.76
6	1,443	60.57	1.65	6	13.58
7	1,145	55.81	1.80	6	14.67
8	977	53.94	1.88	7	16.58
9	876	50.57	2.03	9	17.35
10	691	51.81	2.08	10	18.23
11	570	48.77	2.22	10	22.46
12	517	45.46	2.28	10	18.76
13	392	50.26	2.36	12	21.68
14	366	53.28	2.11	12	16.39
≥ 15	2,216	42.24	2.89	28	21.21
Any	16,565	62.69	1.79	28	12.90

% Samp. 1st: % of transactions where the advert bought was sampled first.
Mean/Max: average/maximum sampling order of advert bought.
% Keep Samp.: % of transactions where the advert bought was not sampled last.

The utility cost of search frictions. We end this section by quantifying the utility loss incurred by consumers because of search frictions. In partial equilibrium, these frictions have two negative consequences for consumers. First, consumers may not purchase the best available advert, as they have not sampled this advert. Secondly, they have to pay a drawing cost every time they sample an advert.[35]

We show these two costs in Table 10 using our model estimates. The second column shows the share of transactions with a minimum search cost $\underline{s}$ in a specific range among all explained transactions. For each explained transaction, we set s and γ respectively equal to the minimum search cost $\underline{s}$ and to the corresponding minimum MWP $\underline{\gamma}(\underline{s})$. We can then compute, for each transaction, the utility of the advert that was bought and of the best available advert for this transaction, which then gives the utility loss for the consumer derived from not buying the best available advert. This is shown in the third column of Table 10, averaged over transactions with a minimum search cost $\underline{s}$ in a given range. We can also compute the minimum number of draws undertaken before sampling the advert that was bought and multiply it by the search cost $\underline{s}$ to find the minimum sampling cost. The minimum number of draws for transaction i is the number of adverts j that have $r_j > r_i$, hence we do not count the first draw. This cost is shown in column 4 of Table 10. Summing the advert utility loss and the sampling cost gives the overall utility loss due to search frictions. The last column of Table 10 shows the ratio of the cost of drawing over this total utility loss.

As we know from Table 7, more than three quarters of our explained transactions can be rationalised by a perfect-information model. For these transactions the utility loss from search frictions is zero. For the remaining 22% of transactions however, Table 10 shows that the utility loss can be substantial as consumers buy an advert which is on average €6.5 worse (in equivalent utility terms) than the best available advert and they also had to incur a sampling cost of almost €1 to sample the advert they bought (recall that this cost does not include the first draw).[36]

We note that the advert utility loss increases with search cost, as does the cost of drawing, even though the number of draws to reach the bought advert decreases. In short, as consumers face a higher search cost they on average choose worse adverts (relative to the best one) and search less. The last column of Table 10 shows that the relative sampling costs decreases with search costs, going from 19% to 6.6% of the overall utility loss.

[35] In full equilibrium, a change in search frictions would also change the prices posted by sellers and thus change the distribution of prices and advert characteristics.

[36] Utility is measured in Euros because the coefficient of price is normalised to -1 in equation (1).

Table 10: Utility cost of search frictions

	Share	Advert utility loss	Min. sampling cost	Share of sampling cost
Search cost				
$\underline{s} = 0$	77.66%	0.00	0.00	-
$\underline{s} > 0$	22.34%	6.52	0.98	12.8%
$0 < \underline{s} < 1$	5.70%	1.81	0.37	19.0%
$1 \leq \underline{s} < 2$	3.31%	5.23	1.31	19.6%
$2 \leq \underline{s} < 3$	3.46%	7.36	1.41	13.4%
$3 \leq \underline{s} < 4$	4.72%	6.98	0.83	6.50%
$4 \leq \underline{s} < 5$	2.34%	8.64	1.05	6.62%
$5 \leq \underline{s} < 6$	0.99%	12.32	1.65	9.10%
$6 \leq \underline{s}$	1.83%	14.88	1.43	6.59%

Note: Search cost $\underline{s}$ (1^{st} column) is the smallest search cost s such that the transaction can be explained by our model. Share (2^{nd} column) is the share of transactions with a given search cost among explained transactions. Advert utility loss (3^{rd} column) is the average difference in utility between the best available advert and the advert bought. The (average) minimum utility cost of draws required to reach the sold advert is in the 4^{th} column and the last column shows the average of the ratio between the utility cost of drawing and the total utility loss from search frictions.

7 Extension: allowing for unobserved product characteristics

So far our analysis has assumed that all advert characteristics entering consumers' utility were observed by the econometrician. As we discussed in section 3.2, there might be at least two sources of unobserved heterogeneity between adverts as sellers may offer additional shipping services and their text description may offer relevant information to the consumer (for instance more details on a used item condition). More generally, even if we think that unobserved heterogeneity is not likely to be a major issue on the PriceMinister website, it is useful to discuss such an extension of our approach so that it can be applied to other settings where econometricians may not observe all the relevant information. In this section, we first present the extended model and the resulting consumers' strategy. We then adapt our multi-step identification and estimation approach to this modified setting. Lastly, we discuss the new results and compare them with our benchmark case.

7.1 An extended model with unobserved advert characteristics

We now add an advert characteristic, denoted as ε, to the model outlined in section 2. This characteristic is not observed in the dataset but is observed by the consumer (upon paying a search cost) and thus is potentially relevant to his purchasing decision. As in section 4.1, we use a linear projection (which we detail below) of observed advert characteristics onto a scalar index x and let the (homogenous) parameter β characterise this projection. Hence an advert is now characterised by a scalar index x, a price p and an unobserved scalar variable ε. The utility (now denoted $\overline{u}$) that a consumer with preferences γ derives from buying an advert is given by:

$$\overline{u}(p, x, \varepsilon, \gamma) = \gamma(x + \varepsilon) - p. \tag{15}$$

As before, the non-price advert characteristics (x, ε) are not observed by the consumer unless he samples the advert, at a cost s. We assume that the unobserved component is i.i.d. across adverts and consumers, normally distributed with zero mean and a standard error σ. The c.d.f. of this distribution will be denoted as Φ_σ. What follows does not hinge on this normality assumption but our estimation method requires a parametric assumption on this distribution.

In this modified environment, we can still use the sequential directed search result from Weitzman (1979) to characterise the optimal search-and-purchase strategy since we have assumed ε to be independent of x and p. The reservation utility $\overline{r}(p,s,\gamma)$ is now defined in a similar fashion as in (2), the difference being that we now have to take expectations with respect to both x and ε:

$$s = \int\int_{\overline{u}(p,x,\varepsilon,\gamma)>\overline{r}(p,s,\gamma)} \left[\overline{u}(p,x,\varepsilon,\gamma) - \overline{r}(p,s,\gamma)\right] dF_{X|P=p}(x) d\Phi_\sigma(\varepsilon). \tag{16}$$

We can then get a closed-form expression of the reservation utility:

$$\overline{r}(p,s,\gamma) = \gamma \cdot \Omega_p^{-1}\left(\frac{s}{\gamma}\right) - p, \quad \text{where} \quad \Omega_p(X) = E_\varepsilon\left[\psi_p(X-\varepsilon)\right]. \tag{17}$$

Replacing the u and r functions with $\overline{u}$ and $\overline{r}$ respectively, we can apply the same reasoning as in the model described in section 2 and identify the (s,γ) parameter sets with the inequalities in (3)-(4) and (9)-(10). However, these inequalities now involve the ε terms which are not observed by the econometrician and thus have to be integrated out. For a given transaction i and a value of (s,γ), we can no longer claim that the model is rejected or not with certainty but we can compute the likelihood of this transaction given values for the MWP γ, the search cost s, the standard error σ, the projection β and the beliefs ψ. For notational convenience we will write the likelihood of transaction i as $\ell_i(s,\gamma)$ thus making implicit its dependence on σ, β and ψ, which will be estimated in preliminary steps described in the next section.

We now briefly derive the likelihood of a transaction and present our estimation results in this new setting. We start with the case $\gamma = 0$ (and $s \geq 0$) i.e. the consumer only cares about price. The likelihood of transaction i is then equal to zero if the sold advert i is not the cheapest, and it is equal to 1 divided by the number of adverts at the lowest price otherwise (since the consumer will randomise between these adverts).

If $s = 0$ and $\gamma > 0$, the likelihood of transaction i is the probability that $\overline{u}_i$ is larger than the utility of any alternative advert $\overline{u}_j$.[37] We can then write the likelihood of transaction i as follows:

$$\ell_i(s=0,\gamma) = \int_\varepsilon \left[\prod_{j\neq i} \Phi_\sigma\left(\frac{p_j - p_i}{\gamma} + x_i - x_j + \varepsilon\right)\right] d\Phi_\sigma(\varepsilon). \tag{18}$$

Lastly, for the case where $s > 0$ and $\gamma > 0$, we need to consider the inequalities (9)-(10), where $\overline{u}$ and $\overline{r}$ replace u and r respectively, and to integrate over all ε's (that is ε_i for the sold advert and the ε_j's for the other adverts). After some algebra we get the following expression of the likelihood for the case $s > 0$ and $\gamma > 0$:

$$\begin{aligned}\ell_i(s,\gamma) &= \left[1 - \Phi_\sigma\left(\frac{\overline{r}_i + p_i}{\gamma} - x_i\right)\right] \cdot \prod_{\overline{r}_j > \overline{r}_i} \Phi_\sigma\left(\frac{\overline{r}_i + p_j}{\gamma} - x_j\right) \\ &+ \int^{\frac{\overline{r}_i + p_i}{\gamma} - x_i} \left[\prod_{j\neq i, \varepsilon < \frac{\overline{r}_j + p_i}{\gamma} - x_i} \Phi_\sigma\left(\frac{p_j - p_i}{\gamma} + x_i - x_j + \varepsilon\right)\right] d\Phi_\sigma(\varepsilon).\end{aligned} \tag{19}$$

[37] As before, we sometimes denote the utility offered by advert i as $\overline{u}_i$ instead of $\overline{u}\,(p_i, x_i, \varepsilon_i, s, \gamma)$.

7.2 Estimating the extended model

The previous section showed that for any transaction i and parameter values (s, γ), we can compute the extended model's likelihood $\ell_i(s, \gamma)$. To this end, we need the projection β to compute the scalar hedonic index x, the standard error of the shock σ and the consumer belief ψ. These are obtained in a couple of first steps prior to computing the likelihood.

Scalar hedonic index and standard error of the unobserved characteristic. The first step consists in estimating both the linear projection of advert characteristics onto a scalar index x and the standard deviation σ of the unobserved characteristic ε. Once again, let β be the parameter characterising the projection. We are thus looking for an estimate of (β, σ).

As we did in section 4.1 for the benchmark model, we focus on transactions with two adverts, for computational feasibility. Consider such a transaction where advert i is sold and j is not. Let $L_i(0)$ be the maximum over γ of the perfect-information likelihood $\ell_i(0, \gamma)$ given in (18). This quantity is thus the highest probability that transaction i can be explained by the model with no search cost. If i is strictly cheaper than j then this probability is simply equal to 1. If $p_i > p_j$, since $\ell_i(0, \gamma)$ is increasing with γ we have $L_i(0) = \ell_i(0, \overline{\gamma})$, where $\overline{\gamma}$ is the highest possible value for the MWP.[38] Lastly, if $p_i = p_j$, then $L_i(0)$ is the maximum between $1/2$ (corresponding to the case $\gamma = 0$ where the consumer randomises) and $\ell_i(0, \overline{\gamma})$ (corresponding to the case $\gamma > 0$).

Note that the quantities $L_i(0)$ defined above depend on the value of (β, σ). However they do not depend on the consumers' beliefs ψ. Our estimate of (β, σ) is the value that maximises $\sum_{J=2} \ln [L_i(0)]$, where the sum is taken over transactions with two adverts. This is not exactly a sample likelihood because $L_i(0)$ is not an individual likelihood but the highest likelihood of transaction i within the perfect-information model ($s = 0$). The motivation for this estimate of (β, σ) is the following: we maximise the fit of the perfect-information model on two-advert transactions. This is slightly different from what we did for the benchmark model, where we maximised the fit whilst keeping search costs to a minimum. We had to take a different approach for the extended model for computational feasibility.

Consumer beliefs. The estimation of consumers' beliefs ψ is close to what we did for the benchmark model in section 4.2. The ψ function is estimated non-parametrically then it is integrated over ε to get Ω (see equation (17)). We then approximate the inverse $\Omega_p^{-1}(y)$ with a cubic polynomial in p and y, while imposing the partial derivatives with respect to p and y to be positive and negative respectively. The reservation utilities can then be computed from these 'smooth' beliefs using (17).

Computing the individual likelihoods. Once we have the (β, σ) parameters and the Ω_p belief functions, we can compute the individual likelihood $\ell_i(s, \gamma)$ of each transaction for any pair (s, γ). We thus browse a

[38] To prove this first note that $\ell_i(0,0) = 0$ if $p_i > p_j$. Then, with two adverts, we note that the difference between the two adverts' unobserved terms $\varepsilon_i - \varepsilon_j$ follows a centered normal with variance $2\sigma^2$ and we can thus rewrite $\ell_i(0, \gamma)$ as $1 - \Phi_{\sigma\sqrt{2}}\left(\frac{p_i - p_j}{\gamma} + x_j - x_i\right)$ when $\gamma > 0$. In the case $p_i > p_j$, this is increasing in γ.

grid of values for these two parameters and compute and record the individual likelihood for each transaction, using (18)-(19).[39]

Inferring on the joint distribution of s and γ. The previous estimation steps produce for each transaction and each parameter value a likelihood which we can interpret as a posterior probability. In this last step, we produce the prior probabilities for the pair (s, γ).

To this end, we make the assumption that the joint distribution of these parameters is discrete and that its support is a subset of the grid we used for estimation. Ideally, we would estimate both the support points and the prior probabilities using, say, an EM algorithm. This would be difficult in practice as the M-step is costly to implement and we would need a large number of classes to cover all the $s = 0, \gamma > 0$ cases. Instead, we choose to set the support of (s, γ) within our known grid and then estimate the prior probabilities i.e. the weight of each point on that grid. This approach follows Train (2008) and considers a discrete distribution for (s, γ) with fixed points where shares are treated as parameters. The estimation algorithm is easy and fast to implement.[40]

We start with a grid of $C = 45 \times 47 = 2,115$ fixed points for the pair (s, γ). The likelihood for each grid point and each transaction i is computed using (18)-(19), and denoted as ℓ_{ic}. The initial values of the shares (prior probabilities) are denoted as $\mu_c^{(0)}$, $c = 1, ..., C$. Then, for any iteration $k \geq 0$, we loop over the next two steps:

- calculate the posterior probabilities $q_i^{(k)} = \mu_c^{(k)} \ell_{ic} \Big/ \sum_{c'} \mu_{c'}^{(k)} \ell_{ic'}$.
- update the share: $\mu_c^{(k+1)}$ as the empirical population average of $q_i^{(k)}$.

We repeat these two steps until convergence and obtain estimates $\widehat{\mu}_c$ of the shares for each value of (s, γ) on our grid. Note that this is essentially an EM algorithm where the updating of the support points in the M step has been omitted because we work with fixed grid points.

7.3 Results

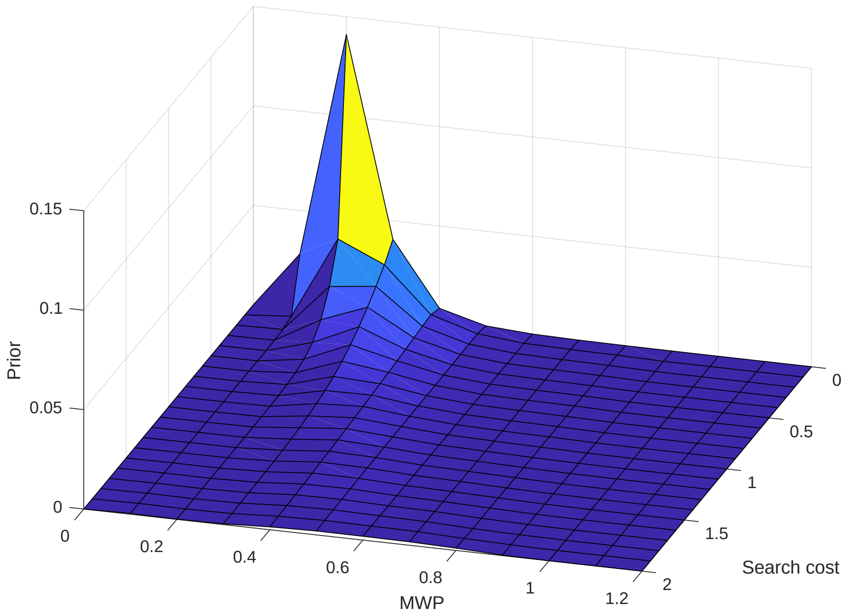
We show in Figure 5 the estimated shares of each point on the grid of search cost s and MWP γ. Since most of the weight is on small values, we have zoomed in on results for $\gamma \leq 1.2$ and $s \leq 2$. We also show in Table 11 the grid points with the highest shares, i.e. all the priors estimated strictly above 1% (in decreasing order) together with their associated values of s and γ.

The first comment we can make about Figure 5 and Table 11 is that there is a mass (14%) with a search cost of zero and a strictly positive MWP (0.2). This means that there is a large share of transactions that are most likely to be explained by a perfect-information model, allowing for consumers to have a 'taste' for non-price advert characteristics.

[39] Given that this step is quite costly, we use a coarser grid than with our benchmark estimation. For s we use a step of 0.1 between 0.1 and 3, a step of 0.2 between 3 and 5 and then each integer between 5 and 10. For γ we use the same grid as for s and add grid points at 15 and 20.

[40] Train (2008) showed that the main drawback of this approach is that summary statistics may be sensitive to the range of the fixed points. Our results should not be much affected by this however as we are mainly interested in showing the distribution of (s, γ) near 0, in order to show evidence of strictly positive search costs and MWPs.

Figure 5: Estimated prior probabilities for each (s, γ) pair



The second main comment is that we find evidence of strictly positive search costs. In fact, summing all the prior probabilities for pairs (s, γ) where $s > 0$, we obtain 77.69%. Our results also suggest some heterogeneity in search costs. Indeed, estimated priors with a search cost $0 < s \leq 1$ add up to 38.5% and those with a search cost above 3 add up to 19.4%. Besides, we find evidence of a strictly positive MWP for the non-price advert characteristics. Its value is concentrated below 1: values of γ strictly below 1 (resp. below 2) have a cumulated share of 97.7% (resp. 99.92%).

Overall, the estimates from this extended model with unobserved advert characteristics are very much in line with those of our benchmark model. In either case, we find evidence of strictly positive search costs and MWP for non-price attributes.[41] We also find evidence of heterogeneity in search costs and of large search costs for a small number of transactions. We thus conclude from this exercise that the results we obtained with our benchmark model are robust to the introduction of an unobserved advert characteristic. Moreover, this section has shown how to take the sequential directed search model to the data whilst allowing for unobserved characteristics and keeping a flexible modelling of the heterogeneity in preferences and search costs.

[41] We note that the estimation results put almost no weight on $\gamma = 0$. Given our specification of the utility function (15), the likelihood of a transaction given $\gamma = 0$ is equal to 0 if the cheapest advert was not sold. This is not at odds with the results from our benchmark model however. Indeed, we can see in Figure 2 that a distribution that puts no weight on $\gamma = 0$ fits easily between the estimated bounds.

Table 11: Grid (s, γ) points with prior probabilities strictly larger than 1%

γ	0.2	0.2	0.3	0.3	0.1	0.3	0.2
s	0	0.1	0	0.1	0	0.2	0.2
Prior (%)	14.01	4.34	4.07	3.31	2.81	2.73	2.46
γ	0.3	0.3	0.3	0.2	0.8	0.4	0.3
s	0.3	0.4	0.5	0.3	4.6	0.1	0.6
Prior (%)	2.20	1.73	1.34	1.31	1.09	1.05	1.02

8 Conclusion

We have conducted a structural analysis of consumer preferences and search costs on the Internet allowing for flexible heterogeneity along these two dimensions. Estimating sets of demand-side parameters from a sequential directed search model, we can account for a wide range of search patterns, where the sampling order depends on both the preferences of consumers and on their search cost. In particular consumers may not necessarily sample adverts in a monotonic (decreasing or increasing) price order. We show robust evidence on the role played by search costs and individual preferences in online transactions. In particular, we show that a flexible modelling of unobserved heterogeneity is important. Indeed, while the majority of transactions can be explained by a perfect information model, we find that a substantial share of purchases (more than a fifth) must have been made by consumers facing positive, sometimes high, search costs.

As far as we know, this paper is innovative on the methodological front by taking an empirical approach inspired by the revealed preferences literature to a search model. First, we derive inequalities characterising the sets of preference and search cost parameters which follow from the optimal search-and-purchase strategy. Then we show how to feasibly take these inequalities to the data, thanks to a first-stage projection that reduces the dimensionality of the problem. This approach allows us to account for flexible heterogeneity in preferences and search costs as well as for interactions between consumer preferences and search behaviour. This latter feature exploits the potential of the Weitzman (1979) model to describe a varied set of ordered search behaviours. Also, our approach does not require data on sampling behaviour and could thus be implemented in a wide range of applications where only data on transactions and adverts are available.

A possible direction for future extension of this framework would consist in closing our model and thus solving for sellers' optimal price posting strategy given consumers' search behaviour. The characterisation of the identified sets that we derived from the optimal search-and-purchase strategies could be used to compute a seller's probability of making a transaction at a given price conditionally on the distribution of preferences and search costs. This distribution could be parametrised whilst ensuring that it fits within the bounds found in our partial equilibrium analysis. We could then assess the respective roles of the heterogeneity in consumer preferences, search costs and seller characteristics in the substantial price dispersion observed in the Internet and documented in this paper. Closing a consumer search model *à la* Weitzman (1979) however is not straightforward and issues revolving around heterogeneity on either or both sides of the market and on the information available to each seller prior to posting their advert need to be investigated. The paper by Anderson and Renault (2016) provides a recent survey of the consumer search literature and, focusing more on the framework of Weitzman (1979), the contribution by Armstrong (2016) provides a useful starting point for a structural equilibrium analysis of an ordered search model.

References

Analytis, P., A. Kothiyal, and K. Katsikopoulos (2014): "Multi-Attribute Utility Models as Cognitive Search Engines," *Judgement and Decision Making*, 9, 403–19.

Anderson, S. and R. Renault (1999): "Pricing, Product Diversity, and Search Costs: a Bertrand-Chamberlin-Diamond Model with Consumer Search," *Rand Journal of Economics*, 30, 719–35.

——— (2016): "Firm Pricing with Consumer Search," *CEPR Discussion Paper*, DP11723.

Armstrong, M. (2016): "Ordered Consumer Search," *CEPR Discussion Paper*, DP11566.

Baye, M., J. Morgan, and P. Scholten (2004): "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site," *Journal of Industrial Economics*, 52, 463–496.

Blow, L., M. Browning, and I. Crawford (2008): "Revealed Preference Analysis of Characteristics Models," *Review of Economic Studies*, 75, 371–89.

Bronars, S. G. (1987): "The power of nonparametric tests of preference maximization," *Econometrica*, 693–698.

Burdett, K. and K. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, 51, 955–69.

Chen, Y. and S. Yao (2016): "Sequential Search with Refinement: Model and Application with Click-stream Data," *Management Science*, forthcoming.

Cherchye, L. Crawford, I., B. De Rock, and F. Vermeulen (2009): "The Revealed Preference Approach to Demand," in *Quantifying Consumer Preferences*, ed. by D. Slottje, Bingley, UK: Emerald Books.

Cosaert, S. and T. Demuynck (2014): "Revealed preference theory for finite choice sets," *Economic Theory*, 56, 1–32.

De Los Santos, B., A. Hortaçsu, and M. Wildenbesst (2012): "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102, 2955–80.

——— (2017): "Search with Learning for differentiated products: evidence from E-commerce," *Journal of Economics and Business Statistics*, 35, 626–41.

Diamond, P. (1971): "A model of Price Adjustment," *Journal of Economic Theory*, 3, 156–68.

Dinerstein, M., L. Einav, J. Levin, and N. Sundaresan (2014): "Consumer Price Search and Platform Design in Internet Commerce," mimeo.

Hadar, J. and W. R. Russell (1969): "Rules for ordering uncertain prospects," *The American Economic Review*, 59, 25–34.

Hong, H. and M. Shum (2006): "Using price distributions to estimate search costs," *RAND Journal of Economics*, 37, 257–275.

Honka, E. and P. Chintagunta (2016): "Simultaneous or Sequential? Search Strategies in the U.S. Auto Insurance Industry," *Marketing Science*, 36, 21–42.

Hortaçsu, A. and C. Syverson (2004): "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds," *Quarterly Journal of Economics*, 119, 403–56.

Janssen, M. and J. Moraga-González (2004): "Strategic Pricing, Consumer Search and the Number of Firms," *Review of Economic Studies*, 71, 1089–1118.

Kim, J., P. Albuquerque, and B. Bronnenberg (2010): "Online Demand Under Limited Consumer Search," *Marketing Science*, 29, 1001–1023.

——— (2016): "The Probit Choice Model Under Sequential Search With An Application To Online Retailing," *Management Science*, forthcoming.

Koulayev, S. (2014): "Estimating demand in online search markets, with application to hotel bookings," *RAND Journal of Economics*, 45, 553–75.

Moraga-González, J. and V. Petrikaite (2013): "Search Costs, Demand-Side Economies and the Incentives to Merge under Bertrand Competition," *Rand Journal of Economics*, 44, 391–424.

Moraga-González, J., Z. Sándor, and M. Wildenbeest (2013): "Semi-Nonparametric Estimation of Consumer Search Costs," *Journal of Applied Econometrics*, 28, 1205–23.

Moraga-González, J. and M. Wildenbeest (2008): "Maximum Likelihood Estimation of Search Costs," *European Economic Review*, 52, 820–48.

Morgan, P. and R. Manning (1985): "Optimal Search," *Econometrica*, 53, 923–44.

Stahl II, D. (1989): "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review*, 79, 700–12.

Stigler, G. (1961): "The Economics of Information," *Journal of Political Economy*, 69, 213–25.

Stoye, J. (2010): "Partial Identification of Spread Parameters," *Quantitative Economics*, 1, 323–57.

Train, K. (2008): "EM Algorithms for nonparametric estimation of mixing distributions," *Journal of Choice Modelling*, 1, 40–69.

Ursu, R. (2017): "The Power of Rankings," *Management Science*, forthcoming.

Weitzman, M. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47, 641–54.

Wolinsky, A. (1986): "True Monopolistic Competition as a Result of Imperfect Information," *Quarterly Journal of Economics*, 101, 493–511.

Zhou, J. (2011): "Ordered Search in Differentiated Markets," *International Journal of Industrial Organization*, 29, 253–262.

APPENDIX

A Formal presentation of the sequential search problem.

Transposing the model of Weitzman (1979) into our setting, we consider a consumer faced with a set $\mathcal{J} = \{1, ..., J\}$ of adverts and with a utility of not buying anything of u_0. Once the consumer has sampled a subset $\mathcal{S}$ of adverts, he has the choice of sampling one of the remaining adverts in the set $\overline{\mathcal{S}} = \mathcal{J} \setminus S$ or settling with the best utility he has found so far: $\tilde{u} = \max\{u_0, \max_{j\in\mathcal{S}}\{u_j\}\}$. Each advert j offers a utilty u_j with a cumulative distribution function of $G_j(.) = F_{X|P}(\cdot|p_j)$. The dynamic problem faced by the consumer can be formalised as follows (this is our adaptation of equation (2) in Weitzman, 1979). The consumer's asset value of having a set of $\overline{S}$ adverts yet to sample and a best utility found so far of $\tilde{u}$ is:

$$U(\overline{\mathcal{S}}, \tilde{u}) = max\{\tilde{u}, max_{j\in\overline{\mathcal{S}}}\{W_j\left(\overline{\mathcal{S}}, \tilde{u}\right)\}\}$$

where $W_j\left(\overline{\mathcal{S}}, \tilde{u}\right)$ is the value of sampling advert j when one has found a highest utility level of $\tilde{u}$ and has a set $\overline{\mathcal{S}}$ of adverts yet to sample:

$$W_j\left(\overline{\mathcal{S}}, \tilde{u}\right) = -s + \delta\left[G_j(\tilde{u})U\left(\overline{\mathcal{S}} \setminus \{j\}, \tilde{u}\right) + \int_{\tilde{u}} U\left(\overline{\mathcal{S}} \setminus \{j\}, u_j\right) dG_j(u_j)\right],$$

where s is the consumer's search cost and δ is the discount factor since Weitzman (1979) considers that there may be a waiting time to discover the contents of each of Pandora's boxes. In our case however, we consider that the time elapsed when finding out about the characteristics of an advert is very small so we will set the discount factor to 1. The reservation threshold r_j for advert j is the utility level $\tilde{u}$ such that the consumer is indifferent between sampling j or settling with r_j. Formally:

$$s = \int_{r_j} (u_j - r_j)dG(u_j)$$

As shown in Weitzman (1979) the optimal strategy boils down to opening the box with the highest reservation threshold r_j among the adverts in $\overline{\mathcal{S}}$ and to carry on doing so until $\tilde{u}$ exceeds the highest r in the remaining adverts. Formally, when the consumer carries on sampling, the next advert sampled is advert j^* where:

$$j^* = argmax_{j\in\overline{\mathcal{S}}}\{W_j\} = argmax_{j\in\overline{\mathcal{S}}}\{r_j\}$$

B Proof of Proposition 2.1

The set of values of (s, γ) which rationalise transaction i, S_i, is the intersection of the sets S_{ij} that can explain that advert i has been bought whilst advert j was available but not chosen. Proposition 2.1 in section 2.3 states that when $s > 0$ and $\gamma > 0$, the pair (s, γ) will not belong to S_{ij} if and only if conditions (5)-(6) are verified. We now prove that proposition.

The tree in Figure A-1 summarises all possible cases where $r_i \neq r_j$ and the resulting sampling and buying decisions. In two cases, when $u_i \geq u_j$, we state that the consumer buys item i with a strictly positive probability. This follows from our assumption that when both i and j have been sampled, the consumer will always choose i if $u_i > u_j$ and he will randomise with equal probabilities between the two adverts if $u_i = u_j$. In either case, i will be bought with a strictly positive probability so we cannot reject that the consumer would choose i.

From the choices illustrated in Figure A-1, we reject that the consumer buys item i (i.e. the probability that the consumer chooses i when j is available is zero) if and only if:

$$\{(r_i < r_j) \wedge [(u_j \geq r_i) \vee ((u_j < r_i) \wedge (u_i < u_j))]\} \vee \{(r_i > r_j) \wedge (u_i < r_j) \wedge (u_j > u_i)\}$$

Taking intersections with $(u_i \geq r_i) \vee (u_i < r_i)$ the above statement is equivalent to:

$$\{(u_i \geq r_i) \wedge (r_i < r_j) \wedge (r_i \leq u_j)\} \vee \{(u_i < r_i) \wedge (r_i < r_j) \wedge (r_i \leq u_j)\} \vee ...$$
$$\{(u_i < r_i) \wedge (r_i < r_j) \wedge (u_j < r_i) \wedge (u_i < u_j)\} \vee \{(u_i < r_i) \wedge (r_j < r_i) \wedge (u_i < r_j) \wedge (u_i < u_j)\}$$

$$\iff \{(u_i \geq r_i) \wedge (r_i < r_j) \wedge (r_i \leq u_j)\} \vee ...$$
$$(u_i < r_i) \wedge \big\{(r_i < r_j) \wedge \big[(r_i \leq u_j) \vee \big((u_j < r_i) \wedge (u_i < u_j)\big)\big] \vee [(r_j < r_i) \wedge (u_i < r_j) \wedge (u_i < u_j)]\big\}$$

$$\iff \{(u_i \geq r_i) \wedge (r_i < r_j) \wedge (r_i \leq u_j)\} \vee ...$$
$$(u_i < r_i) \wedge \{[(r_i < r_j) \wedge (u_i < u_j)] \vee [(r_j < r_i) \wedge (u_i < r_j) \wedge (u_i < u_j)]\}$$

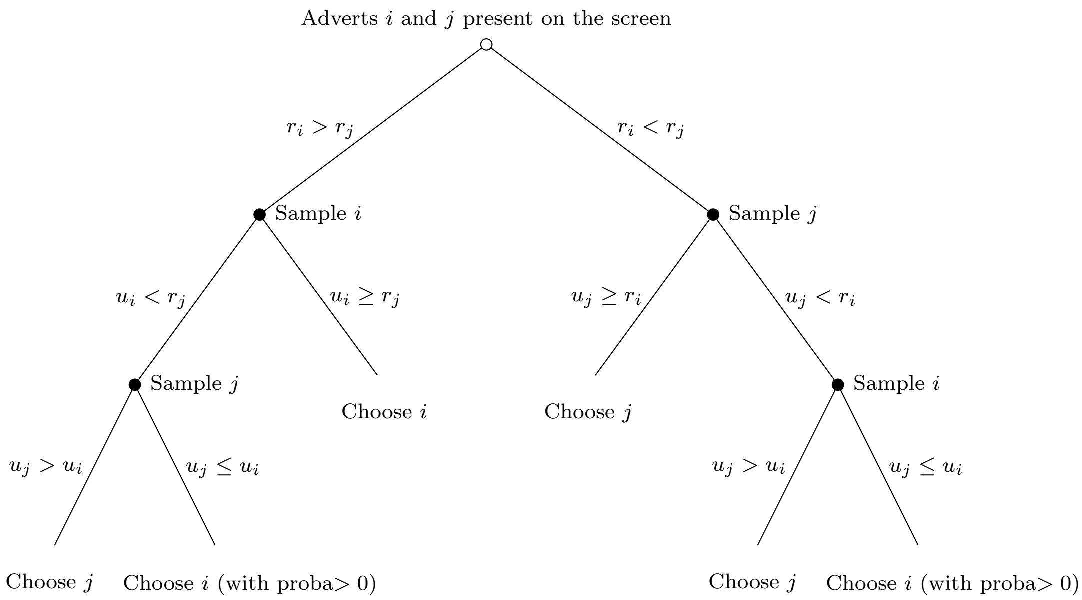
$$\iff \{(u_i \geq r_i) \wedge (r_i < r_j) \wedge (r_i \leq u_j)\} \vee \{(u_i < r_i) \wedge (u_i < u_j) \wedge (u_i < r_j)\}$$

To complete the proof, we need to consider the case where $r_i = r_j$. Then the consumer will sample i or j first with equal probabilities, i.e. he will go down the right or the left branch at the top node of the tree in Figure A-1 with equal probabilities. In this case, the probability that the consumer chooses i is equal to 0 if and only if:

$$(r_i = r_j) \wedge (u_i < r_j) \wedge (u_i < u_j) \wedge \{(r_i \leq u_j) \vee [(u_j < r_i) \wedge (u_i < u_j)]\}$$
$$\iff (r_i = r_j) \wedge (u_i < r_j) \wedge (u_i < u_j)$$

Combining this with what we found above for the case $r_i \neq r_j$, we get the necessary and sufficient conditions (5) and (6) in Proposition 2.1.

Figure A-1: Choice between adverts i and j



C Construction of the sample

The administrative data we obtained from PriceMinister consist of two tables. The first table is a record of *all* transactions that took place on the website until December 2008. For each transaction, we observe, among other things, the seller id, the product id, the advert id (not the buyer's), the price, the exact date when the transaction was initiated/completed, the seller's status (professional or individual) and the feedback given by the consumer for this specific transaction. With this information we can thus compute each seller's size (number of completed transactions so far) and reputation (average feedback received so far) at any given date. We observe a seller's status unless she has no transactions, initiated or completed (i.e. only appears in the table of adverts). We assume that sellers who never appear in the transaction table are private individuals (expecting professional sellers to have at least one contact with a buyer).

The second table contains all the adverts posted on the website, with information on advert id, seller id, product id, the condition of the good (new, as new, etc.), list price, the precise date when the advert was posted and whether the advert is still active at the data extraction date.

We combine these two tables to produce a dataset that, for any transaction in a given time period and product category (in our benchmark case, CDs during the last quarter of 2007), provides information on the adverts for the exact same product available at the time when the transaction took place. To this end, we had to solve two problems.

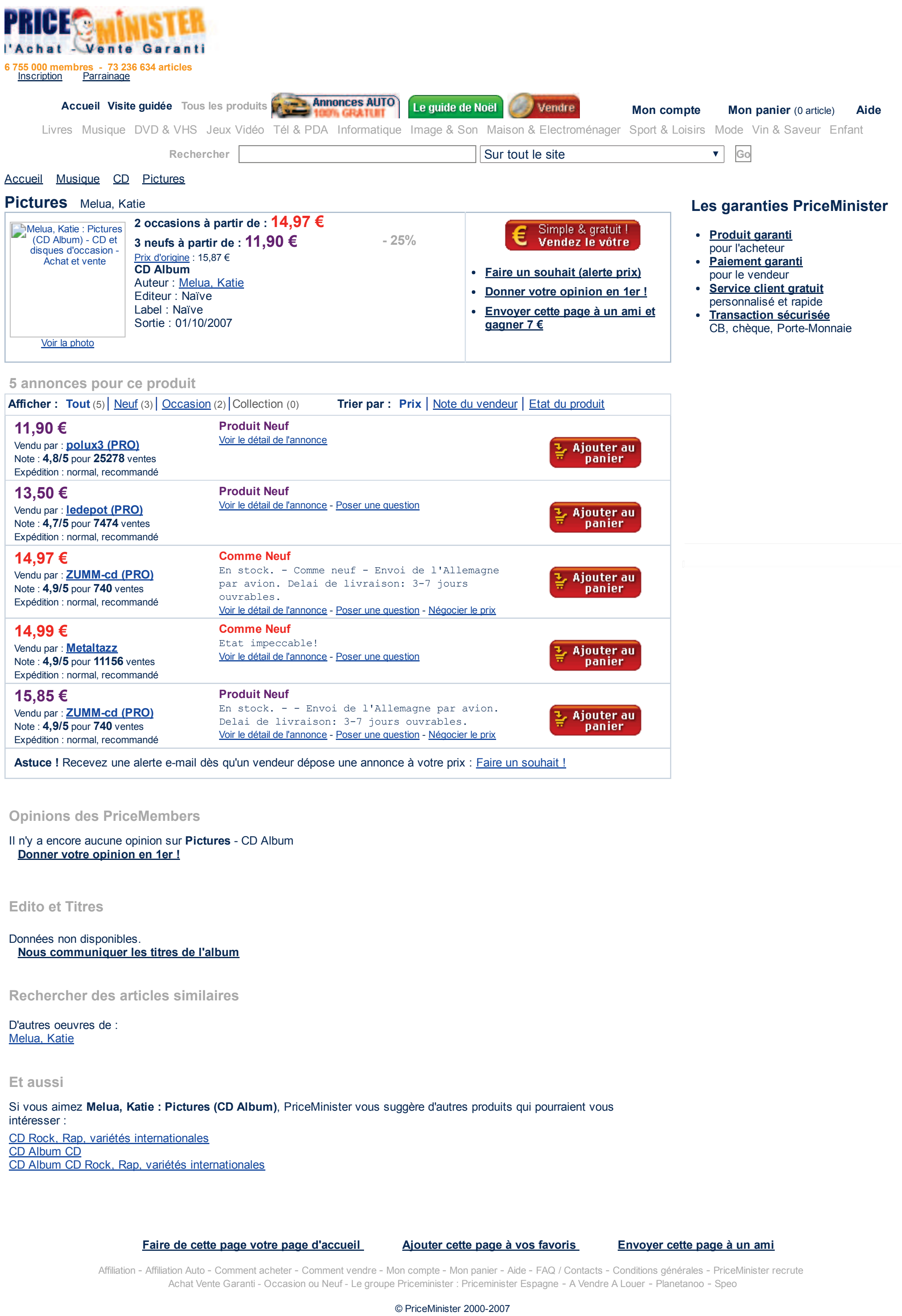
The first issue is that the match on advert id between the transaction and the advert tables is not perfect. For a small proportion of CDs (which is the product category we are interested in), there are transactions for which the advert id is not found in the advert table. We cannot just get rid of these transactions because the advert that was bought on this occasion may have been available to consumers in other transactions. After some investigation, we have reached the conclusion that this problem, which again only concerns a small minority of CDs, may be caused by adverts that are sold very quickly, within a few hours of being posted and thus appear in the transaction table but have not yet been included in the advert table. To make sure that these adverts do not interfere with other transactions for the same product, we drop all observations for a given product during the day when such a mismatched transaction takes place. A more drastic solution would consist in leaving out of the sample all the CDs for which this mismatch takes place, at any time. This would however take out the bestselling products and severely decrease the number of transactions. A previous version of the paper used that correction and the results were qualitatively similar to those shown in this version. Hence, we chose to keep as many products as possible, including the bestselling ones, and to apply the first, less drastic, solution to solve the mismatch problem.

The second problem pertains to adverts' end date i.e. when adverts disappear from the website. Our advert data includes each advert's creation date and activity status in December 2008. There are thus adverts for which the end date is not directly observed and we had to construct these dates based on a few assumptions. If an advert is still active in December 2008, we assume that it has been active since its creation. If an advert is no longer active at the extraction date and has led to at least one transaction, we assume that it was closed (taken off the screen) right after its last observed transaction. This is to reflect the fact that the seller ran out of stocks after the last transaction. Indeed, most sellers are individuals who probably have only one copy to sell. Professionals, on the other hand, are unlikely to deactivate an advert for a product that sells if stocks are still positive. The last and most difficult case concerns adverts that are inactive in December 2008 but were never sold. We assume that these adverts were taken off by sellers after a given time period which we compute as follows. We consider the distribution of durations between a CD advert's first transaction and its creation (conditionally on being sold at least once). We use the 95% quantile of this distribution as an assumed cutoff time at which these adverts are taken off the website. In section G we show that the main results on fit and the importance of search costs are not changed when assuming that these adverts are not closed during our observation period.

D Screenshot of an advert page

We show in Figure A-2 a screenshot for a CD in October 2007. The photograph of the album cover cannot be displayed here but it is the same for all adverts on that page. There are five adverts for this product, sorted by increasing price order. The price is in big digits on the left and the additional information on the product condition's and the seller's reputation, size and status is in smaller print. The "negotiate the price" feature was a recent feature at the time and was very rarely used so we ignored it. The source of unobserved advert heterogeneity may come from the different type of shipping offered by the seller (the shipping fee is set by PriceMinister) and from the text comment. In this example however, all sellers offer the same shipping options. The extended model estimated in section 7 aimed at capturing these unobserved attributes. Lastly, one may claim that since all the relevant information is already on the screen, consumers can observe it at no cost. This is allowed in our benchmark estimation as a value of zero for the search cost is included in our grid. Still, we find that 22% of transactions cannot be explained by the perfect-information model.

Figure A-2: Screenshot of an advert page for a CD in October 2007



E Additional descriptive statistics

In this section, we show the distribution of non-price characteristics, first unconditionally on prices and then conditionally on the price being in one of four quartiles. Table 12 shows the distributions in the population of adverts and Table 13 shows the same statistics in the population of transactions.

The first panel of each table shows that the ratio of professional to individual sellers is about one third if we do not condition on prices and that this ratio tends to increase as we move up price quartiles.

The distribution of product condition is shown in the second panel. The condition “fair”, which is attached to only a few adverts, is merged with the next best one, “good”. This table shows that trade on the PriceMinister website mostly involves second-hand items. We note that although the second-hand product condition is self-reported, sellers do not systematically post the best condition. In fact, for about 40% of adverts, the product condition is advertised as less than “as new”. The distribution of product condition varies slightly across price quartiles but always displays substantial variance.

The third and fourth panels of Tables 12-13 show several quantiles of the distribution of seller size whilst conditioning on seller status. Note that these size distributions are taken in the population of adverts and transactions respectively, not in the population of sellers. We observe a substantial dispersion of seller size, whether we control for price quartile or not. As expected, professional sellers can complete far more transactions than individual sellers, which adds to the heterogeneity in advert characteristics.

The last panel of 12 (resp. 13) shows the distribution of seller reputation among adverts (resp. transactions). Recall that each transaction is completed once feedback, i.e. an integer between 1 and 5, is given by the buyer. The seller’s average feedback rounded to the first decimal is then shown on all his live adverts. We note that in most cases, an advert or a transaction is associated with a seller’s reputation higher than 4.5. There is substantial dispersion between 4.5 and 5, even if we focus on a price quartile.

Table 12: Distribution of characteristics among adverts, overall and by price quartile

	All adverts	1$^{\text{st}}$ quart.	2$^{\text{nd}}$ quart.	3$^{\text{rd}}$ quart.	4$^{\text{th}}$ quart.
Seller status: (%)					
- *Professional*	30.39	12.52	22.35	38.03	48.44
- *Individual*	69.61	87.48	77.65	61.97	51.56
Product condition: (%)					
- *Good*	10.2	19.04	11.3	5.97	4.54
- *Very Good*	29.41	36.64	33.5	27.61	20.02
- *As New*	42.27	38.65	43.75	44.07	42.62
- *New*	18.12	5.67	11.46	22.35	32.81
Seller size (Pro):					
- 10% *Quantile*	216	181	247	226	211
- 25% *Quantile*	777	592	874	765	780
- *Median*	3,192	2,610	4,064	4,441	2,900
- 75% *Quantile*	12,066	13,703	89,167	89,401	5,631
- 90% *Quantile*	96,871	99,906	100,633	101,222	23,762
- *Mean*	23,045	25,222	32,158	32,959	10,740
Seller size (Ind):					
- 10% *Quantile*	5	4	5	6	5
- 25% *Quantile*	27	23	25	31	31
- *Median*	117	96	109	142	156
- 75% *Quantile*	435	344	420	509	562
- 90% *Quantile*	1,247	912	1,243	1,508	1,491
- *Mean*	459	377	444	513	557
Seller reputation: (%)					
- $r \leq 4$	1.38	1.64	1.54	1.19	1.13
- $4.1 \leq r \leq 4.5$	14.61	11.07	15.44	20.89	11.16
- $r = 4.6$	6.9	7.36	7.08	7.13	6.04
- $r = 4.7$	11.93	13.28	12.17	10.49	11.76
- $r = 4.8$	27.29	26.4	26.73	25.32	30.64
- $r = 4.9$	27.04	25.52	25.38	25.94	31.25
- $r = 5$	7.65	10.62	8.29	6.32	5.41

Table 13: Distribution of characteristics among transactions, overall and by price quartile

	All transactions	1st quart.	2nd quart.	3rd quart.	4th quart.
Seller status: (%)					
- *Professional*	37.92	20.71	34.61	53.14	51.58
- *Individual*	62.08	79.29	65.39	46.86	48.42
Product condition: (%)					
- *Good*	7.65	14.64	7.11	3.32	2.15
- *Very Good*	21.81	32.17	23.43	13.57	12.7
- *As New*	38.76	37.76	41.47	35.76	41.21
- *New*	31.78	15.43	28	47.35	43.94
Seller size (Pro):					
- 10% *Quantile*	473	340	514	994	300
- 25% *Quantile*	2,532	1,378	3,245	10,362	1,122
- *Median*	86,344	12,653	90,965	94,237	5,362
- 75% *Quantile*	104,442	102,302	105,798	108,507	88,556
- 90% *Quantile*	119,755	119,179	120,832	123,322	106,845
- *Mean*	59,690	50,583	64,728	74,405	35,666
Seller size (Ind):					
- 10% *Quantile*	8	7	8	9	9
- 25% *Quantile*	34	30	33	43	44
- *Median*	135	109	134	180	205
- 75% *Quantile*	460	383	453	563	637
- 90% *Quantile*	1,198	935	1,195	1,467	1,572
- *Mean*	461	400	459	516	573
Seller reputation: (%)					
- $r \leq 4$	0.54	0.79	0.59	0.33	0.29
- $4.1 \leq r \leq 4.5$	25.69	15.63	25.54	40.94	20.01
- $r = 4.6$	4.77	5.96	5.32	3.79	3.18
- $r = 4.7$	9.45	11.58	9.26	7.03	9.69
- $r = 4.8$	23.24	25.11	23.2	19.02	26.75
- $r = 4.9$	26.43	28.2	26.16	21.93	30.99
- $r = 5$	8.03	10.29	8.09	5.72	7.39

F Robustness check: 'raw' vs 'smooth' consumer beliefs

We have chosen to estimate sets of parameters and not to impose restrictions on the distribution of heterogeneity. While our modelling approach has been flexible with respect to the distribution of the parameters (s, γ), we approximated the ψ_p^{-1} functions to force the beliefs on the distribution $F_{X|P}$ to be smooth and monotonic in p (see section 4.2). This seems desirable to represent the beliefs of consumers regarding advert 'quality' given price. In this section, we show the estimation results obtained when using the 'raw' beliefs $\widetilde{\psi}$ computed directly from the data using equation (12), and compare them with our benchmark results.

Table 14 shows the fit of the model. Comparing these results with the benchmark ones in Table 6, we note that the fit on all transactions remains very good (92.6%). Hence the fact that our model can explain most transactions taking place on the platform does not hinge on the constraints we have put on consumers' beliefs. This remains true if we focus on transactions with a 'better' advert, as the fit is still 73% for this case. As expected, the fit on transactions with no 'better' advert is still almost perfect (98.6%).

Table 14: Pass rate (%) among transactions with J adverts - Raw beliefs

	All transactions		Transactions with no 'better' advert		Transactions with $\geq$ 1 'better' advert	
# adverts	Freq.	% Pass	Freq.	% Pass	Freq.	% Pass
2	19,248	99.09	17,180	100.00	2,068	92.00
3	13,234	97.34	11,100	99.69	2,134	86.26
4	9,597	95.13	7,579	99.18	2,018	81.19
5	6,975	92.90	5,309	98.53	1,666	77.56
6	5,415	90.71	3,975	97.94	1,440	73.13
7	4,106	88.75	3,045	96.99	1,061	69.14
8	3,244	88.01	2,308	97.77	936	67.50
9	2,645	85.60	1,866	96.43	779	63.99
10	2,174	85.51	1,515	96.47	659	63.89
11	1,707	85.00	1,173	96.47	534	62.30
12	1,472	84.17	999	96.13	473	63.72
13	1,192	82.38	794	95.30	398	60.94
14	1,063	79.30	698	96.66	365	50.99
$\geq$ 15	5,681	79.30	3,502	94.26	2,179	53.39
Any	77,753	92.60	61,043	98.62	16,710	73.09

Freq: number of transactions in each category.

Pass: proportion (in %) of transactions that can be explained - $S_i \neq \{\emptyset\}$.

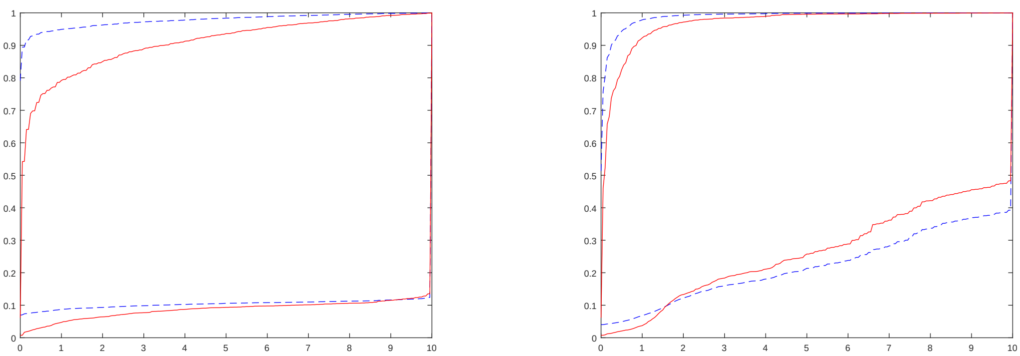
We now show the lower and upper bounds on the c.d.f. of the search cost s and MWP γ in Figure A-3. For each parameter, we show bounds on the c.d.f. for all transactions and also for transactions with a 'better' advert. Comparing these figures with Figures 1-2 where 'smoothed' beliefs were used, we have two main comments.

First, for any belief specification ('smooth' or 'raw') or group of transactions (all or with a 'better' advert), we have empirical evidence of strictly positive MWPs and search costs. Indeed, there is a fraction of transactions that cannot be explained by a perfect-information model. In other words, our finding on the presence of search costs on this Internet platform does not depend on the 'smooth' restrictions we have put on consumers' beliefs.

The second comment is that the bounds are wider when using 'raw' beliefs (Figure A-3) than when using 'smooth' ones (Figures 1-2). This was expected as tightening the bounds comes at the cost of imposing some restrictions on the beliefs. What we show in this section however is that these restrictions have no qualitative impact on the results.

Figure A-3: Bounds on the c.d.f. of search cost s and MWP γ

Search cost *MWP*



Vertical axis: lower and upper bounds - Horizontal axis: s *(left) and* γ *(right)*

Dashed: All explained transactions - Solid: Explained transactions with a 'better' advert

The main difference between the two sets of results relates to the predictions on the magnitude of search costs. For explained transactions with 'better' adverts, the median of $\underline{s}$ is at least 2.9 when using 'smooth' beliefs (see Figure 1) whereas the estimate using 'raw' beliefs is strictly positive but small (see Figure A-3). These results are not contradictory but indicate that the model with 'smooth' beliefs requires relatively high search costs to fit the data better. Importantly, this does not mean that the median search cost is above 2 Euros. Indeed, Table 7 shows that for 77.66% of explained transactions, the search cost can be 0. If we focus however on the smaller group of transactions with a 'better' advert and which are thus likely to have strictly positive search costs, then the search cost can be quite high.

This is intuitive as transactions with 'better' adverts are often such that $p_i > p_j$ (the advert sold is strictly more expensive than the 'better' alternative). For this transaction to be explained by the model, it must be that i is drawn and j is not. If i is more expensive, consumers must expect a better x in order to sample i first. This will not always be the case if $\psi_p(x)$ does not increase with p. Hence the 'raw' beliefs may not allow the model to explain some of these transactions. This will not happen however with the 'smooth' specification where it is always possible to find a MWP large enough so that i is drawn before j, since beliefs regarding the quality offered by an item priced p_i stochastically dominate beliefs regarding the quality of a cheaper item. We also need a high enough search cost to rationalise that, once i is drawn, j is not sampled.

What we take from this section is that our results are robust to relaxing the assumption we made in the main analysis that consumers believe that quality increases smoothly with price at all points of the distribution of qualities. We run further robustness checks with alternative samples relating to a different date or a different product in Appendix G. These checks also support the robustness of our findings.

G Additional results.

Advert closure date. As we explain in section C, to construct our sample we need to impute the closing date of adverts who are never sold and are inactive at the extraction date (December 2008). As a benchmark, we have assumed that these adverts were closed after a duration which corresponds to the 95% quantile of durations until first sale for adverts which are sold at least once. In this section, we consider an alternative and set these adverts' closing date to the extraction date i.e. they are assumed not to be closed during our observation period (third quarter of 2007). The main estimation results are shown in Table 15, with the benchmark sample in the first column and the alternative closing date sample in the second column. We can see that the fit is left unchanged, whether we look at transactions with or without a 'better' advert. The proportion of transactions requiring a strictly positive search cost is also stable (22.3% vs 24.6%). The distribution of the minimum search cost (respectively MWP) is slightly shifted to the right (respectively left) which, after further investigation, seems to come from the first stage of our estimation. Indeed, since the first-stage criterion is based on a pass rate for transactions with two adverts, and since dropping out/keeping adverts due to the closing date may remove some of these transactions from the sample, the resulting estimate of β may be sensitive to the closing date. Still, this variability does not affect our main results on the role of search costs and preferences.

Alternative quarter or product. The results in this paper have pertained to CD transactions during the third quarter of 2007. In this section we now show that our results still hold when considering other time periods or another product category (DVD). We use two alternative samples: CD transactions during the second quarter of 2007 and DVD transactions during the third quarter of 2007. We apply the same selection criteria for products (catalog price between 10 and 25€) and transactions (no advert has a price below 1 or above 20€) as in our benchmark sample (see section 3.3). We present in Table 15 the main estimation results obtained when using each of these samples.

The main results on search costs and preferences using these alternative samples are similar to those we obtain in our benchmark case (shown in the first column of Table 15). In the fourth column, we see that consumer preferences and search costs also play an important role in DVD transactions. Even though the fit is slightly lower than for CDs, we note that our model can explain 87% of all transactions and 62% of transactions where a 'better' advert was available. Results also show that some consumers are willing to pay substantially more for better advert characteristics and can face relatively high search costs. In particular, 29% of the explained transactions for DVDs are not consistent with a perfect-information model.

Power of the approach. While our estimation results show that our model can rationalise a very high fraction of the transactions we observe in our data, a remaining question is whether our approach would also accommodate other behaviours, i.e. if our approach is able to discriminate between "rational" datasets and less rational ones. Bronars (1987) proposes a simple way of assessing power by evaluating the probability that the model rejects a set of (simulated) data where individuals make choices at random. We have thus simulated a dataset where consumers, faced with the same adverts as in our benchmark estimation, buy one advert at random. We have then applied our estimation method to this data set and we show the results in the last column of Table 15. When estimated on this simulated dataset, our model fit is much worse than with our real data (74.7% vs 95.4%), which reassures us about its ability to discriminate. Moreover, with random choices, strictly positive search costs are needed for 36% of the transactions explained, as opposed to 22% with the real data; and search costs greater than 4 euros are needed for 28% of the transactions explained (as opposed to 4.7%). In short, when estimated on random choice data, the fit of the model deteriorates substantially and higher, less realistic, search costs are needed to explain the transactions that can be rationalised.

Note that is difficult to compare our measure of power to usual examples in the literature on consumer choice for two reasons. First, as choice is discrete in our environment and many transactions occur with only 2 or 3 adverts on the screen, the counterfactual dataset that we produce with random choices coincides with actual choices in 27% of cases, which means that we could not reject more than 73% of the random transactions on average. Secondly, power in revealed preference tests usually refers to rejecting non-rational behaviours in a perfect-information setting. Our search model however can accept behaviours that would be rejected by a rational perfect-information model since we allow for search frictions. The rate of rejection in the random sample should thus not be seen as the sole indicator of power as it should be put in parallel with the distribution of search costs. Hence, the fact that the distribution of search costs substantially shifts to the right when going from the data to the random sample indicates that our approach can help to discriminate between rational (with imperfect information) and random consumer behaviour.

Table 15: Results using alternative samples

	Estimation sample				
	CD 2007Q3	CD (alt) 2007Q3	CD 2007Q2	DVD 2007Q3	CD (rand) 2007Q3
pass rate (%) among transactions					
- any	95.36	94.79	95.76	87.23	74.70
- with no 'better' advert	99.80	99.84	99.78	99.62	99.33
- with ≥ 1 'better' advert	79.16	79.71	81.37	61.78	50.57
share (%) of explained transactions with $\underline{s} > 0$					
- any	22.34	24.58	23.08	28.97	36.40
- with no 'better' advert	6.66	5.90	6.85	9.34	5.22
- with ≥ 1 'better' advert	94.57	94.38	94.23	93.99	96.39
share (%) of explained transactions with:					
- $\underline{s} = 0$	77.66	75.42	76.92	71.03	63.61
- $0 < \underline{s} \leq 0.1$	2.59	1.56	2.04	2.35	1.60
- $0.1 < \underline{s} \leq 0.5$	1.77	1.69	1.43	2.16	1.42
- $0.5 < \underline{s} \leq 1$	1.51	1.57	1.48	2.24	1.12
- $1 < \underline{s} \leq 2$	3.35	2.60	3.41	4.81	2.18
- $2 < \underline{s} \leq 3$	3.40	2.39	3.44	4.15	1.11
- $3 < \underline{s} \leq 4$	5.01	2.21	2.76	6.58	1.27
- $4 < \underline{s} \leq 5$	1.96	6.65	2.05	1.88	23.78
- $5 < \underline{s}$	2.75	5.91	6.46	4.80	3.91
share (%) of explained transactions with:					
- $\underline{\gamma} = 0$	49.87	50.22	49.40	45.91	37.84
- $0 < \underline{\gamma} \leq 0.1$	21.44	26.99	25.92	15.04	43.83
- $0.1 < \underline{\gamma} \leq 0.5$	11.66	16.12	14.72	11.16	6.25
- $0.5 < \underline{\gamma} \leq 1$	7.44	1.88	8.04	4.82	2.70
- $1 < \underline{\gamma} \leq 1.5$	8.17	1.83	1.50	5.29	1.27
- $1.5 < \underline{\gamma} \leq 2$	1.20	1.17	0.34	6.06	0.65
- $2 < \underline{\gamma}$	0.22	1.79	0.10	11.72	7.47